UNIVERSIDAD DE GRANADA

FACULTAD DE CIENCIAS

DEPARTAMENTO DE ÓPTICA

PROGRAMA DE DOCTORADO EN FÍSICA Y CIENCIAS DEL ESPACIO

Hyperspectral image processing for material identification in historical documents and artworks

Tesis Doctoral

Ana Belén López Baldomero

Directores

Eva María Valero Benito

Miguel Ángel Martínez Domingo

Granada, 2025

Hyperspectral image processing for material identification in historical documents and artworks

Memoria presentada por la Graduada en Óptica y Optometría,

Máster en Optometría Clínica y Óptica Avanzada, y

Máster en Física: Radiaciones, Nanotecnología, Partículas y Astrofísica

Ana Belén López Baldomero

para optar al Grado de Doctora por la Universidad de Granada con

Mención de Doctorado Internacional

Fdo: Ana Belén López Baldomero

Granada, marzo 2025

Editor: Universidad de Granada. Tesis Doctorales
Autor: Ana Belén López Baldomero
ISBN: 978-84-1195-820-2
URI: https://hdl.handle.net/10481/105339

La doctoranda / *The doctoral candidate* **Ana Belén López Baldomero** y los directores de la tesis / *and the thesis supervisors* **Eva María Valero Benito** y **Miguel Ángel Martínez Domingo**:

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor/s and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / *Place and date*: Granada, marzo 2025

Directores de la Tesis:

Fdo: Eva María Valero Benito y
Miguel Ángel Martínez Domingo

Doctoranda:

Fdo: Ana Belén López Baldomero

A mis padres, mi hermano y Antonio

Agradecimientos (Acknowledgments)

Una vez alguien me dijo que una tesis doctoral era como una carrera de obstáculos y, aunque es cierto que estos tres últimos años han estado llenos de subidas y bajadas emocionales, siempre he sido consciente de lo agradecida que debía estar. Primero, por haber tenido la oportunidad de hacer la tesis. Segundo, por, además, haber podido trabajar en un tema que me ha motivado tanto desde el inicio y que a día de hoy me sigue motivando a seguir en la investigación. Pero, sobre todo y lo más importante, estoy muy agradecida de la gente que me he encontrado en este camino y también de aquella que me ha apoyado en todo momento.

En primer lugar, quiero dar las gracias a mis directores de tesis, Eva Valero y Miguel Ángel Martínez. Creo que cualquier cosa que pueda escribir aquí no es suficiente para agradecer todo lo que me habéis ayudado en este tiempo. Os admiro de corazón, no sólo en lo laboral, sino también en lo personal. Eva, gracias por amar

tanto lo que haces y saber transmitir esa motivación a los demás, por tu inagotable energía, por tu comprensión, tu amabilidad, y por ser una fuente de inspiración. Migue, gracias por guiarme y apoyarme tan bien, incluso en los momentos más difíciles, por esa dedicación, por estar siempre disponible y, ante todo, gracias por ser un gran amigo. Gracias, a ambos, por enseñarme tanto y por confiar en mí. Espero que esto sea sólo el pequeño comienzo de un largo camino juntos.

Al *Color Imaging Lab*, por acogerme y hacerme sentir como en casa desde el primer momento. A Javi, Juan Luis, Javier, Fran, Ramón y Yannick. Gracias a todos por vuestra ayuda para hacer esto posible.

Thank you, Prof. Sony George, from the *Norwegian University of Science and Technology (NTNU)*, for welcoming me, for your constant support in moving the research forward, and for sharing your knowledge. I truly appreciate all your valuable help and guidance. To my colleagues at *NTNU*, thank you for making me feel at home during those three months, I miss you all. A special thanks to Marco Buzzelli from the *University of Milano–Bicocca* for all the fruitful online meeting discussions.

Al Departamento de Pintura y al Grupo de Química Analítica y Ciencias de la Vida de la UGR, en especial a Rosario Blanc, Ana López y Anna Reichert, por aportar su conocimiento imprescindible para esta tesis. A Carolina Cardell, del Departamento de Mineralogía, por su valiosa colaboración y por permitirnos utilizar su instrumento. A David Torres Ibáñez, Director del Archivo de la Real Chancillería de Granada, y Eva Martín López, Directora del Archivo Histórico Provincial de Granada, por facilitarnos el acceso a sus archivos.

A todos los compañeros y compañeras del Departamento de Óptica de la UGR, gracias por integrarme y hacerme sentir una profesora más. Al director del Departamento de Óptica, Rafael Huertas, por su buena disposición en todo momento y por darme un espacio donde poder trabajar. A todos los profesores con los que he

tenido el gusto de compartir asignatura, gracias por enseñarme tanto. A Manuel Rubiño, Carlos Salas y Carolina Ortiz, por darme la oportunidad de comenzar en la investigación.

A mis amigos del departamento, María, Pilar, Javi, Sonia, Aitor, Francesco, Miriam, Juancho, Luis y Rafa. Gracias por todo, por las comidas en la biblioteca, las tapas de los viernes, el esquí, el perreo, el senderismo, los días de playa... Pero sobre todo, gracias por estar ahí y convertiros casi en una familia para mi.

A mis amigas y amigos de toda (o media) vida, Andrea, Brenda, Erika, Bea, Jonny y Masi. Por ser un gran apoyo y confiar siempre en mí. Por estar en los buenos momentos y, sobre todo, en los malos. Porque el tiempo siga sin pasar por nuestra amistad.

A mi madre, mi padre y mi hermano, gracias por hacerme ser quien soy a día de hoy, porque sin vosotros no habría llegado hasta aquí. Por ese cariño incondicional, no tengo palabras para agradecer todo lo que habéis hecho por mi. Os quiero muchísimo. A toda mi familia, gracias por el apoyo, y en especial a mis abuelas. Esto es por vosotras.

A Antonio, por ser un pilar fundamental en mi vida. Por levantarme, aguantarme, y cuidarme cuando más lo necesitaba. Por quererme tanto y tan bien. Por todo lo que me has enseñado, y lo más importante, por hacerme mejor persona y ayudarme a valorar las cosas importantes de la vida. Te quiero.

A todos y todas, GRACIAS.

Contents

Abbreviations and Acronyms

λ: Wavelength

AZ: Azurite

BB: Bones Black

BBONE: Bones black

BIVOR: Ivory black

BLAMP: Lampblack

BRDF: Bidirectional Reflectance Distribution Function

BVINE: Vine black

CC: Carbon-containing ink

cGFC: complement of the Goodness-of-Fit coefficient

CH: Cultural Heritage

CN: Cinnabar

CNN: Convolutional Neural Network

CP: Carbon-based ink pure

CS: Conservation Science

DE00: CIEDE2000 color difference

DeGu: DeepGun library

DL: Deep Learning

DT: Decision Trees

EM: Endmember

EMX: Endmember number X (where X = 1,2,...)

FIPPI: Fast Iterative Pixel Purity Index

FN: False Negatives

FORS: Fiber Optic Reflectance Spectroscopy

FP: False Positives

FTIR: Fourier Transform Infrared spectroscopy

GFC: Goodness-of-Fit Coefficient

GT: Ground Truth

HSI: Hyperspectral Imaging

ICH: Intangible Cultural Heritage

IGILA: Mixture of lampblack with iron gall ink

IGINK: Iron gall ink

IGISE: Mixture of sepia with iron gall ink

IR: Infrared

JM: Jeffries–Matusita

KNN: k-Nearest Neighbors

LDA: Linear Discriminant Analysis

LIP: Logarithmic Image Processing

LL: Lapislazuli

LW: Lead White

MEx: Manual Extraction library

MGP: Pure metallo-gallate ink

ML: Machine Learning

MSI: Multispectral Imaging

MWIR: Middle Wavelength Infrared

NCC: Non-carbon-containing ink

NFDL: NFINDR library

NIR: Near-Infrared

NMF: Nonnegative Matrix Factorization

NN: Neural Network

NS3: Normalized Spectral Similarity Score

NY: Naples Yellow

OX: Objective X (where X = 1,2,...)

P1: Preparation layer

PCA: Principal Component Analysis

PCs: Principal Components

PLS-DA: Partial Least Squares Discriminant Analysis

PPI: Pixel Purity Index

PSD: Probability of Spectral Discrimination

PWSD: Power of Spectral Discrimination

R: Reflectance

Ref.: References

REFL: Reference library

REP: Read earth

RF: Random Forest

RGB: Red, Green, and Blue

RIS: Reflectance Imaging Spectroscopy

RMSE: Root Mean Square Error

RS: Reflectance Spectroscopy

SAM: Spectral Angle Mapper

SCM: Spectral Correlation Mapper

SEM: Scanning Electron Microscopy

SEPIA: Sepia ink

SID: Spectral Information Divergence

SNR: Signal-to-Noise Ratio

SP: Sepia pure ink

SPD: Spectral Power Distribution

SUPP: Support

SVM: Support Vector Machines

SWIR: Short-Wave Infrared

TCH: Tangible Cultural Heritage

TN: True Negatives

TP: True Positives

UNESCO: United Nations Educational, Scientific and Cultural Organization

UV: Ultraviolet

VAF: Variance Accounted For

VIS: Visible

VNIR: Visible and Near Infrared

VPSEM: Variable Pressure Scanning Electron Microscopy

XRD: X-Ray Diffraction

XRF: X-Ray Fluorescence

Abstract

Material identification in historical documents and artworks is essential for their study, conservation, and authentication. Understanding the composition of inks, pigments, and supports enables researchers to address degradation processes, establish a plausible date, and detect alterations. In this context, hyperspectral imaging (HSI) has emerged as a highly promising technique due to its ability to acquire both spatial and spectral information within a wide range from ultraviolet to mid-wavelength infrared, without physical contact with the objects and in less time than conventional methods.

Although spectral reflectance provides clues about chemical composition, factors such as conservation status, similarity of spectra for different materials, or the presence of mixtures can complicate identification. Given the advantages of HSI, there is a need to develop more generalized and automated methodologies for material identification that are independent of specific features of the document or artwork,

such as the number of materials present, the techniques employed by the artist, the time period or aging effects.

This PhD thesis explores and optimizes various hyperspectral image analysis techniques, including spectral similarity metrics, machine learning models and deep learning, as well as spectral unmixing techniques, with the aim of identifying and mapping inks, pigments, and supports in historical documents and artworks.

One of the main contributions is the development of HYPERDOC, a public database of hyperspectral images of historical documents and mock-up samples. It provides spectral information in the visible and near-infrared (VNIR) and short-wave infrared (SWIR) ranges, along with data on the materials present in each pixel.

First, a preliminary study was carried out to perform material identification using spectral similiary metrics and a reference library of materials. Three methods were proposed to determine the most suitable spectral metrics. The integration of the three methods proved effective for identifying inks in mock-ups and historical manuscripts, although further refinement was required.

For this reason, six models were implemented and compared, including five traditional machine learning methods (SVM, KNN, LDA, RF, and PLS-DA) and a deep learning (DL) model, for the classification of three types of ink: pure metallo-gallate (MGP), carbon-containing (CC), and non-carbon-containing (NCC) inks. The results demonstrated that while DL achieved the highest accuracy, SVM provided comparable performance with lower computational demands. Additionally, post-processing techniques and data fusion of VNIR and SWIR spectral ranges contributed to the high classification performance.

To address the problem of mixed materials, a study was performed on a copper plate oil painting with a reduced palette, pre- and post-intervention. Unmixing techniques were used to perform material identification. Three methods for ex-

tracting endmembers (NFINDR, NMF, and manual extraction) were evaluated in the VNIR range. Manual extraction provided the most accurate pigment concentration maps and spectral reconstruction, while NFINDR and NMF correctly identified more pigments by direct comparison between endmembers and a reference library of pigments. Furthermore, using the VNIR and SWIR ranges, linear and nonlinear unmixing models were compared using two hyperspaces: reflectance (R) and -log(R). In the restored painting, the -log(R) hyperspace with NFINDR proved better results, while the R space allowed complete pigment identification merging results from both ranges.

To extend this technique to historical document analysis, unmixing was applied to identify components in ink mixtures using fused VNIR and SWIR data. The results showed that it is difficult to detect certain components within the mixtures, especially when CC ink is present, since the mixture spectra tend to be remarkably similar to the pure CC ink spectra. Additionally, unmixing was used as preprocessing to reduce the influence of the support on ink spectra along the borders of the written strokes, improving classification with an SVM model. Both approaches showed promising results, although further optimization is needed to reach their full potential in historical document analysis.

The findings of this PhD thesis contribute to the advancement of HSI techniques in the study of cultural heritage, particularly in material identification and mapping. The development of a database, optimization of spectral analysis methods, and integration of machine learning models and unmixing techniques has contributed significantly towards reaching more precise and automated material identification in historical documents and artworks.

Resumen

La identificación de materiales en documentos históricos y obras de arte es esencial para su estudio, conservación y autenticación. Conocer la composición de tintas, pigmentos y soportes permite a los investigadores abordar procesos de degradación, estimar la época de la obra y detectar posibles alteraciones. En este contexto, la imagen hiperespectral ha surgido como una técnica muy prometedora debido a su capacidad para adquirir información espacial y espectral en un amplio rango desde el ultravioleta hasta el infrarrojo de onda corta, de forma no invasiva y en menos tiempo que técnicas convencionales.

Aunque la reflectancia espectral ofrece indicios sobre la composición química, factores como el estado de conservación, la similitud de los espectros de diferentes materiales o la mezcla de ellos, pueden dificultar su identificación. Debido a las ventajas que presenta la imagen hiperespectral, es necesario desarrollar metodologías más generalizadas y automatizadas para la identificación de materiales que no de-

pendan de características específicas del documento u obra de arte, como el número de materiales presentes, las técnicas empleadas por el artista, la época o el envejecimiento sufrido.

En esta tesis doctoral se han explorado y optimizado diversas técnicas de análisis de imágenes hiperespectrales, incluyendo métricas de similitud espectral, modelos de aprendizaje automático (*Machine Learning*) y aprendizaje profundo (*Deep Learning*), así como técnicas de *unmixing* o desmezclado espectral, con el objetivo de identificar las tintas, pigmentos y soportes presentes en documentos históricos y en una obra de arte, y determinar su distribución espacial.

Una de las principales contribuciones es el desarrollo de HYPERDOC, una base de datos pública de imágenes hiperespectrales de documentos históricos y muestras modelo. Esta proporciona información espectral en los rangos visible e infrarrojo cercano (VNIR), e infrarrojo de onda corta (SWIR), junto con datos sobre los materiales presentes en cada píxel.

Primero, se llevó a cabo un estudio preliminar para la identificación de materiales utilizando métricas de similitud espectral y una biblioteca de referencia de materiales. Se propusieron tres métodos para determinar las métricas espectrales más adecuadas. La integración de los tres métodos demostró ser eficaz para identificar tintas en muestras modelo y manuscritos históricos, aunque se requería mayor refinamiento.

Es por ello que se implementaron y compararon seis modelos, incluyendo cinco métodos tradicionales de aprendizaje automático (SVM, KNN, LDA, RF y PLS-DA) y un modelo basado en aprendizaje profundo (DL), para la clasificación de tres tipos de tinta: metalogálica pura (MGP), con contenido de carbón (CC) y sin contenido de carbón (NCC). Los resultados demostraron que, si bien DL logró la mayor precisión, SVM obtuvo un rendimiento comparable con menores demandas computacionales.

Además, las técnicas de postprocesamiento y la fusión de datos espectrales VNIR y SWIR contribuyeron al alto rendimiento de la clasificación.

Para abordar el problema de las mezclas de materiales, se llevó a cabo un estudio en una pintura al óleo sobre placa de cobre con una paleta reducida, antes y después de su restauración. Se utilizaron técnicas de *unmixing* para realizar la identificación de materiales. Se evaluaron tres métodos de extracción de *endmembers* o componentes puros (NFINDR, NMF y extracción manual) en el rango VNIR. La extracción manual proporcionó los mapas de concentración de pigmentos más precisos y mejor reconstrucción espectral, mientras que con NFINDR y NMF se identificaron más pigmentos correctamente mediante la comparación directa entre los *endmembers* y una biblioteca de referencia de pigmentos. Por otro lado, utilizando los rangos VNIR y SWIR, se compararon un modelo lineal y no lineal de *unmixing*, utilizando dos espacios espectrales: reflectancia (R) y -log(R). En la pintura restaurada, el espacio -log(R) con NFINDR dio mejores resultados, mientras que en el espacio R se consiguió una identificación completa con la información de ambos rangos.

Para extender esta técnica al análisis de documentos históricos, se aplicaron técnicas de *unmixing* para identificar componentes en mezclas de tintas mediante la fusión de datos VNIR y SWIR. Los resultados mostraron que es difícil detectar ciertos componentes dentro de las mezclas, especialmente cuando está presente la tinta CC, ya que los espectros de mezcla tienden a ser notablemente similares a los espectros de la tinta CC pura. Además, se aplicó el *unmixing* como preprocesamiento para reducir la influencia del soporte en las reflectancias espectrales de las tintas, especialmente en los bordes de los trazos, mejorando la clasificación con un modelo SVM. Ambos enfoques mostraron resultados prometedores, aunque es necesario optimizar su aplicación para alcanzar su máximo potencial en el análisis de documentos históricos.

Los hallazgos de esta tesis contribuyen al avance de las técnicas de imagen hiperespectral en el estudio del patrimonio cultural, especialmente en la identificación y distribución espacial de materiales. El desarrollo de una base de datos, la optimización de los métodos de análisis espectral y la integración de modelos de aprendizaje automático y técnicas de *unmixing* han contribuido a una identificación más precisa y automatizada de materiales en documentos históricos y obras de arte.

CHAPTER 1

Introduction

1.1 Motivation

From the Alhambra of Granada to the historic Albaicín quarter and the art of flamenco, Cultural Heritage (CH) represents a vast legacy passed down through generations. These are just a few examples from the city where I was born, but different expressions of culture can be found all around the world in different societies, shaped by history, geography, and interactions with the natural environment.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines CH as covering both tangible (TCH) and intangible (ICH) elements of cultural significance [1, 2]. TCH refers to physical objects and sites that hold symbolic, historic, artistic, aesthetic, ethnological, anthropological, scientific, or social value. It includes artifacts, monuments, groups of buildings, archaeological findings, and museums. ICH encompasses practices, representations, expressions, knowledge, and skills that are transmitted through generations and continuously evolving. 2003 UNESCO Convention for the Safeguarding of ICH [3] includes here oral traditions and expressions, performing arts, social practices, rituals, festive events, and knowledge related to nature and the universe. This definition is not static, as it may include additional elements depending on local contexts, such as traditional games or culinary traditions.

CH is universally recognized as a fundamental right and a common good for individuals and communities, as articulated in the 2001 UNESCO Declaration on Cultural Diversity [4]. It is not only a repository of human history, but also a cornerstone of cultural identity and collective memory, and a vital resource for the past, present, and future of humanity. Interpretation, management, and preservation of CH depend on the active participation of individuals and communities, making them key stakeholders and beneficiaries of conservation efforts.

Conservation of CH is critical to ensure that its significance endures over time and to preserve the transmission of its values and messages to future generations [1, 5]. However, CH faces numerous threats, including natural factors such as adverse weather conditions, human-induced damage, and socio-political conflicts, which underscore the need for robust strategies to safeguard heritage. In response to all these threats, the 1972 UNESCO Convention concerning the Protection of the World Cultural and Natural Heritage established a framework for global conservation efforts [6], supported by institutions like the International Council on Monuments and Sites (ICOMOS). Yet, 50 years later, these challenges persist, with added pressures from climate change, political and armed conflicts, and globalization.

CH holds value for individuals and society for various reasons, including aesthetic, emotional, and economic considerations. There is an inherent desire to preserve CH and make it accessible so that it can be enjoyed and further studied by others [7], and that is where heritage science plays its role. This multidisciplinary field, which dates back to the 1800s, integrates expertise from chemistry, physics, biology, engineering, art history, and, as explored in this PhD thesis, optics and computer science, among other disciplines, to study the materials, techniques, and histories of CH. Heritage science encompasses conservation science (CS), archaeological science, and building science. Regardless of the specific object of study, the scientific methods and research approaches are often similar.

CS aims to characterize the physical and chemical properties of historical objects, developing innovative methods and technologies to prevent deterioration, evaluate past conservation treatments, and ensure practices that respect the integrity of TCH to support its preservation, protection, and maintenance [8]. This science bridges the scientific and cultural domains. A key focus of CS is material characterization, which involves identifying pigments, dyes, binders, supports, alteration products,

and other components of TCH. These analyses provide different and valuable insights, including:

- Identifying and understanding the original materials and techniques used by artists and craftsmen [9].
- Analyzing and modeling degradation processes over time [10].
- Diagnosing causes of deterioration [10, 11].
- Guiding the selection of appropriate restoration and preservation strategies [11–13].
- Identifying previous restorations [9].
- Enhancing the understanding of artistic and historical contexts [13, 14].
- Supporting hypotheses for dating and determining the geographic origin of artworks and documents [15].
- Determining authorship and detecting falsifications [16].

In general, the identification of the materials used is essential to effectively preserve, conserve, and restore TCH [17]. For this purpose, different analytical techniques can be employed, classified as either invasive or non-invasive methods. In the past, invasive techniques were commonly used, involving direct intervention or alteration of the heritage object through processes such as micro- and macro-sampling, making small incisions, or scraping surfaces for subsequent laboratory analysis. However, these techniques are now generally discouraged due to the potential damage to the piece under study. Consequently, non-invasive techniques have been encouraged and developed, allowing for the examination of objects without altering their integrity, thereby improving both the quantity and quality of the analyses [18]. The adoption of these techniques has been supported financially by various organi-

zations, with the European Union investing over 500 million euros through its Horizon Europe program under the "Culture, Creativity, and Inclusive Society" cluster. Examples of projects in this field include COSCH [19], CHANGE [20], and ChemiNova [21].

In section 2.2, a summary is provided of the most commonly used non-invasive techniques for CH, focusing on portable methods suitable for on-site use in museums, archaeological sites, and archives. The summary outlines their basic operating principles, the information provided, and their limitations. A key constraint of many of these systems is that they are non-imaging or point-based techniques, which restricts their ability to provide spatially resolved material distribution maps across different areas of the artwork. Alternatively, obtaining such maps often requires significant time, as the artwork must be scanned point by point. Additionally, depending on the technique, other challenges may arise. For instance, some methods provide elemental information, which means they are unable to identify organic compounds or differentiate pigments with similar elemental compositions, such as malachite and verdigris.

In this context, Hyperspectral Imaging (HSI) has become increasingly used for the non-destructive analysis of CH, pigments, and artworks, with an almost 83% increase in the number of published papers on the topic over the past decade (source: Web of Science). This technology integrates spectroscopy and digital imaging, allowing for rapid data acquisition with high spectral and spatial resolution, all without direct contact with the sample. HSI captures the reflectance or transmittance spectrum pixelwise, commonly covering a broad range of the electromagnetic spectrum from the Ultraviolet (UV) through the Visible (VIS), Near-Infrared (NIR), and Short-Wave Infrared (SWIR) regions, although it usually requires different sensors for each range.

Figure 1.1: False RGB images from HSI captures of an authentic manuscript (left) and forgery (right) in the visible (R = 645 nm, G = 565 nm, B = 440 nm), and infrared image of the authentic document (R = 1600 nm, G = 1200 nm, B = 1000 nm) (center).

The potential of HSI in CH applications has been demonstrated in numerous studies (more details in section 2.5). As a particular example, a research project was conducted at the University of Granada by an interdisciplinary team, involving experts in restoration and conservation, chemistry, and the Color Imaging Lab, including the supervisors of this thesis. The study focused on Arabic manuscripts from 1454 to 1492 housed in the Archive of the Royal Chancellery of Granada (see Figure 1.1). HSI analysis revealed the intentional use of different inks within these documents, a practice historically linked to the hierarchical roles of scribes, notaries, witnesses, and judges, with the final signature of the Sultan. This can be seen in the central image of Figure 1.1, where different inks exhibit distinct behaviors in the infrared. However, one document from the 16th century (Figure 1.1 right) presented an anomaly: it was written entirely in a single type of ink. This document was used by a man to claim noble status as a 'hidalgo' and, remarkably, his claim was granted at the time, but centuries later, the forgery was discovered [22]. This case highlights

the ability of HSI not only to reveal historical material practices but also to expose hidden narratives within documents, reinforcing its relevance for the study of CH.

Building on these capabilities, this PhD thesis focuses on material identification from HSI data, as the reflectance spectrum is characteristic of the current status of the material under study, allowing for both identification and spatial mapping of pigments, inks, and supports. While pigment analysis in artworks is also explored, the primary focus of this PhD thesis is on historical documents, which have been less studied in this field. Despite the considerable potential of HSI techniques for analyzing artworks and their successful application to numerous individual cases, material identification remains a significant challenge, especially when trying to rely only on HSI information. To address this, there is a need for the development of more generalized, automated techniques and methodologies that require minimal supervision and are independent of factors such as the specific artwork, the number of materials involved, the artist's techniques, the era of creation, and the aging processes the piece has undergone.

1.2 Hypothesis and objectives

The **main hypothesis** of this PhD thesis is that it is possible to identify the materials used in historical documents and artworks (inks, pigments, and supports) through a fully automated method, using a hyperspectral image of the piece as the starting point.

The reflectance spectrum provides clues about the possible composition of materials. However, reliable identification can only be achieved if the technique has been validated with other non-invasive analytical methods that determine the material present in the piece through point measurements. Therefore, if a large number of

validated analyses are available, this data can be incorporated into an algorithm that learns to identify a specific material.

A major challenge is the lack of spectral data for historical documents covering an extensive time period and reflecting material changes over time. This issue is further exacerbated by the difficulty in finding data that have been validated using other techniques to confirm the materials present. Furthermore, there is currently no automated method for identifying and mapping inks in historical documents using HSI, and identifying pigments in mixtures in real artworks continues to be a challenge.

Taking into account all the above, the following objectives are established for the present PhD thesis:

The **main objective** is to develop automatic methods for analyzing hyperspectral images to identify materials in historical documents and artworks. To achieve this, it is necessary to create a spectral database covering a wide range of samples, including mock-ups created with known components and proportions in the laboratory using various supports and recipes, as well as real documents from different sources and periods. Different spectral metrics and classification techniques will be explored and proposed to enable material identification and mapping.

The **specific objectives** of this PhD thesis are defined as follows:

- **Objective 1 (O1).** Generation of mock-ups, including the most common supports used in historical documents, as well as historical inks following traditional recipes representative of different time periods. Artificial aging of selected samples.

- **Objective 2 (O2).** Spectral image capture of the samples from O1, as well as historical documents from various sources (Provincial Historical Archive of

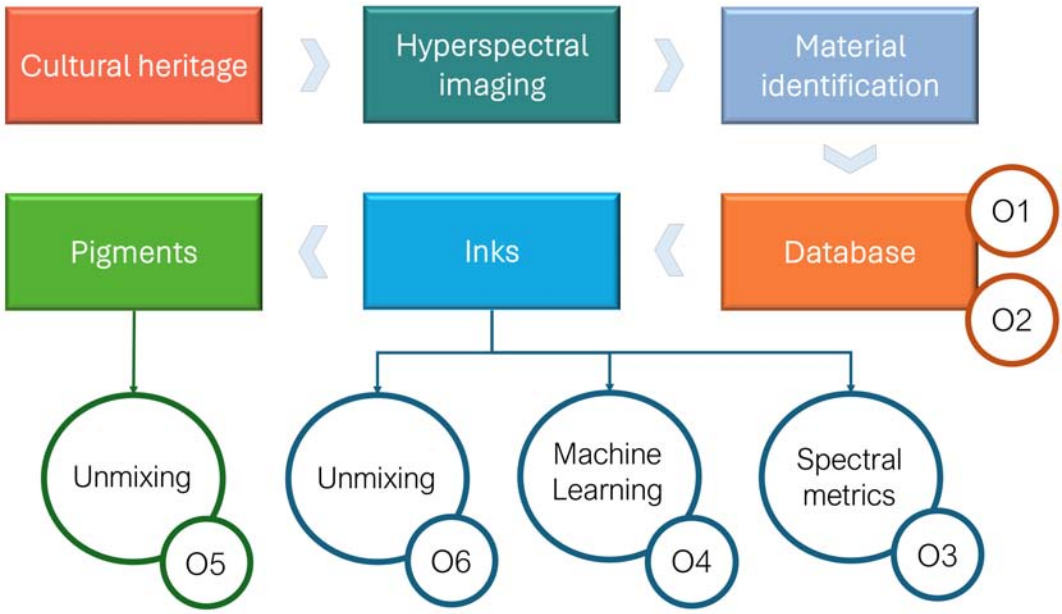

Figure 1.2: Schematic overview of the main topics and specific objectives of this PhD thesis.

Granada, Archive of the Royal Chancellery of Granada). Creation of a public database.

- **Objective 3 (O3).** Implementation of metrics for ink identification through spectral comparison with a reference library, using both mock-ups and historical documents. Selection of the optimal spectral similarity metrics.

- **Objective 4 (O4).** Implementation of Machine Learning (ML) and Deep Learning (DL) algorithms for the identification and mapping of inks in hyperspectral images of historical documents and mock-ups. Optimization of algorithms and selection of the top-performing ones.

- **Objective 5 (O5).** Application and optimization of endmember (EM) extraction and unmixing techniques for identifying pigments and estimating their concentration maps in an artwork.

- **Objective 6 (O6).** Application of unmixing techniques to identify and map components in historical ink mixtures. Implementation as a preprocessing step to enhance ink classification through ML.

1.2 Hipótesis y objetivos

La **hipótesis principal** de esta tesis doctoral es que es posible identificar los materiales presentes en documentos históricos y obras de arte (tintas, pigmentos y soportes) mediante un método completamente automático, utilizando una imagen hiperespectral de la obra como punto de partida.

La reflectancia espectral proporciona pistas sobre la posible composición de los materiales. Sin embargo, una identificación fiable solo puede lograrse si la técnica ha sido validada con otros métodos analíticos no invasivos que determinen el material presente en la obra mediante medidas puntuales. Por lo tanto, si se dispone de un gran número de análisis validados, estos datos pueden incorporarse a un algoritmo que aprenda a identificar un material específico.

Un desafío importante es la falta de datos espectrales de documentos históricos que abarquen un periodo de tiempo extenso y reflejen la evolución en los materiales utilizados. Este problema se ve agravado por la dificultad de encontrar datos que hayan sido validados mediante otras técnicas para confirmar los materiales presentes. Además, actualmente no existe un método automático para identificar y conocer la distribución de diferentes tintas en documentos históricos utilizando imagen hiperespectral, y la identificación de pigmentos en obras de arte reales sigue siendo un reto.

Teniendo en cuenta lo anterior, se establecen los siguientes objetivos para la presente tesis doctoral:

El **objetivo principal** es desarrollar métodos automáticos para analizar imágenes hiperespectrales con el fin de identificar materiales en documentos históricos y obras de arte. Para ello, es necesario crear una base de datos espectral que cubra una amplia gama de materiales, incluyendo muestras modelo creadas en el laboratorio con componentes y proporciones conocidas utilizando diferentes soportes y recetas, así como documentos reales de diversas fuentes y periodos. Se explorarán y propondrán diferentes técnicas de clasificación para identificar los materiales y conocer su distribución en las obras.

Los **objetivos específicos** de esta tesis doctoral se definen de la siguiente manera:

- **Objetivo 1 (O1).** Generación de muestras modelo, incluyendo los soportes más comunes utilizados en documentos históricos, así como tintas históricas siguiendo recetas tradicionales representativas de diferentes periodos. Envejecimiento artificial de muestras seleccionadas.

- **Objetivo 2 (O2).** Captura de imágenes espectrales de las muestras del O1, así como de documentos históricos de diversas fuentes (Archivo Histórico Provincial de Granada, Archivo de la Real Chancillería de Granada). Creación de una base de datos pública.

- **Objetivo 3 (O3).** Implementación de métricas para la identificación de tintas mediante comparación espectral con una biblioteca de referencia, utilizando tanto muestras modelo como manuscritos históricos. Selección de las métricas espectrales óptimas.

- **Objetivo 4 (O4).** Implementación de algoritmos de aprendizaje automático (en inglés, *Machine Learning*) y aprendizaje profundo (*Deep Learning*) para la identificación y mapeo de tintas en imágenes hiperespectrales de documentos

históricos y muestras modelo. Optimización de algoritmos y selección de los más eficientes.

- **Objetivo 5 (O5).** Aplicación y optimización de técnicas de extracción de *end-members* o componentes puros y *unmixing* o desmezclado espectral para la identificación de pigmentos y estimación de sus mapas de concentración en una obra de arte.

- **Objetivo 6 (O6).** Aplicación de técnicas de *unmixing* para identificar y mapear componentes puros en mezclas de tintas históricas. Implementación como preprocesado para mejorar la clasificación a través de algoritmos de aprendizaje automático.

1.3 Structure of the thesis

This PhD thesis is structured according to the specific objectives outlined in section 1.2. After establishing the motivation, hypothesis, and objectives, the chapters are organized as follows:

- **Chapter 2** provides a comprehensive overview of the state of the art in the field, introducing key concepts related to CH and the materials used in artworks and historical documents. It also provides a brief discussion of non-invasive analytical techniques for material identification in this field with a focus on spectral imaging in CH, emphasizing material identification, which is the main objective of this PhD thesis.

- **Chapter 3** introduces the publicly available hyperspectral database, which includes both mock-ups and historical documents from two archives of Granada:

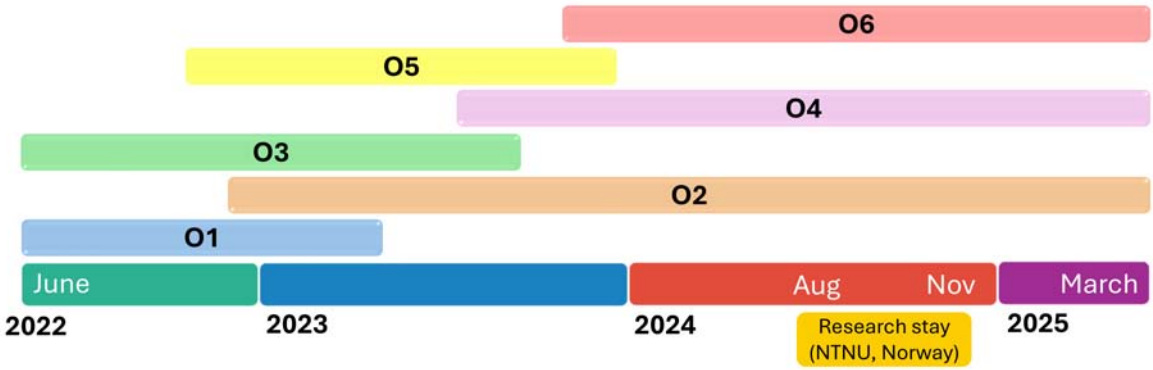

Figure 1.3: Timeline illustrating the progression of the objectives in this PhD thesis.

the Provincial Historical Archive and the Archive of the Royal Chancellery. This chapter addresses specific objectives O1 and O2.

- **Chapter 4** presents three methods for selecting optimal spectral similarity metrics and compares their performance in classifying historical inks in mock-ups and historical documents. This chapter focuses on specific objective O3.

- **Chapter 5** examines and compares the application of traditional ML models and a DL-based approach for ink classification in mock-ups and historical documents. This chapter addresses specific objective O4.

- **Chapter 6** explores various EM extraction methods and unmixing in different hyperspaces and spectral ranges to identify pigments in an oil painting on a copper plate. The optimization of spectral bands for the automatic detection of re-painted areas is also explored. Specific objective O5 is covered in this chapter.

- **Chapter 7** presents the application of spectral unmixing techniques for historical ink analysis with two main purposes: the detection of individual components in ink mixtures, and the use of unmixing as a preprocessing step to

enhance classification in ML models. This chapter focuses on specific objective O6.

- **Chapter 8** summarizes the key findings of this thesis and discusses future research directions.

The research timeline for this PhD thesis is shown in Figure 1.3. The chapters are structured based on the defined objectives, although their development occurred over different periods. During 2024, a research stay was conducted at the Colourlab, Department of Computer Science, Norwegian University of Science and Technology (NTNU), Gjøvik, under the supervision of Prof. Sony George. The results obtained during this research stay are presented in chapter 7.

CHAPTER 2

State of the art

2.1 Materials in historical documents and paintings

In this PhD thesis, the focus will be on the study of two primary sources: historical documents and paintings. Both are valuable components of CH, offering unique insights into the past and enhancing our understanding of human creativity and history. Historical documents serve as repositories of cultural and scientific heritage, capturing knowledge, art and events of human history through manuscripts, archives, and different written or printed works [23–25]. On the other hand, paintings are primarily created for aesthetic expression and appreciation, conveying emotions, ideas, and cultural values. The study of these artworks is of great importance for their preservation and restoration [26–29].

In this context, the term *materials* refers to both the support used (such as paper, parchment, canvas, etc.) and the materials employed for writing, primarily inks, as well as those used for drawing or adding color, such as pigments and dyes. These materials are often combined with other substances, including additives and fillers, which serve to enhance stability and improve specific properties of the inks and pigments. A key component among these materials is the binder, a viscous substance that facilitates cohesion between the coloring particles, and between the particles and the support. Examples of binders include gum arabic, starches, gelatin, albumin, rabbit skin glue and fish glue, among others.

In this PhD thesis, ink is defined as any dark material used for writing; pigment, as an insoluble material, typically inorganic (though not always); and dye, as a soluble material, usually organic and of natural origin. The latter two are primarily used to provide color [18].

In the case of historical documents, some of the most commonly used supports include parchment, which is made from animal skin and was predominantly used

during the Middle Ages, and paper, which varies based on the type of fiber used in its production, including cotton, linen, hemp, and synthetic fibers. On these supports, inks are often applied, as well as color used for illustration purposes, such as in illuminated manuscripts and maps.

Historically, different types of inks have been used across cultures and periods. Their study can reveal much about the sociocultural and technological shifts in historical document production [30]. This makes ink analysis a key tool for codicologists and historians who explore both the content and material composition of manuscripts to gather this information [31]. Inks can be classified into different types based on their origin. This PhD thesis will focus on three groups, chosen for their widespread use throughout history:

1. **Metallo-gallate inks.** These inks primarily consist of two key components: a metallic salt (predominantly ferrous sulfate), and a vegetable tanning agent, typically derived from oak apples in the form of gallotannin extracts, forming ferrous tannate upon mixing. The term "iron gall inks" refers to this category due to its high iron sulfate content [32]. Widely used for centuries, they were valued for their durability, particularly in medieval Europe [30]. Their identification is crucial for selecting optimal conservation strategies, as their composition can cause corrosion of the support material, a phenomenon known as "iron-gall ink burn" [10].

2. **Sepia ink.** This ink has a deep brown color, almost black in full strength. It is obtained from the ink sac of the cuttlefish *Sepia officinalis*. Less stable than carbon-based inks, it is particularly sensitive to light exposure [33, 34].

3. **Carbon-based inks.** The earliest writing ink is thought to have been a carbon-based ink derived from lamp black or other form of charcoal. These inks are

obtained by burning oil or other materials (such as fruit stones, bones, or wood) in controlled conditions with a limited supply of air and mixing them with a binder dissolved in a water-soluble medium [34]. They are very resistant compared with sepia and iron gall inks, being almost insensitive to acids, light exposure, alkalis, and chemicals. They degrade only when the binder loses its mechanical properties.

These three groups of inks could be applied individually or mixed between them to enhance their properties. For example, freshly applied iron gall inks are light brown in color and darken as they oxidize and form ferric tannate. To compensate for this, a small amount of black ink, usually lamp black (carbon-based ink), was often added. The study of mixed inks has received little attention in both scholarly and material research, as their identification and acknowledgment of their significance is a recent development. Medieval Arabic ink recipes suggest that mixed inks held a significant role in the Islamic world [35,36], and mixtures of carbon ink and iron gall ink have been found since ancient times [13, 34], so these types of inks should be given their due consideration.

In paintings, a complex layering of materials is often employed (see Figure 2.1). These layers can generally be categorized into: support (e.g., canvas, wood, metal, paper, etc.), preparatory layer, typically composed of gesso (a mixture of calcium carbonate and glue), which serves a dual function: mechanical (by enhancing the adhesion of the subsequent layers to the support and smoothing out any imperfections in the support itself), and aesthetic, as its base color influences the final appearance of the painting. This is followed by underdrawings, usually created with carbon-based materials such as charcoal or graphite, which serve as a preliminary sketch or guide for the composition. The pictorial layer is where the pigment particles, often in powdered form, are suspended in a binding medium (such as linseed oil for oil painting

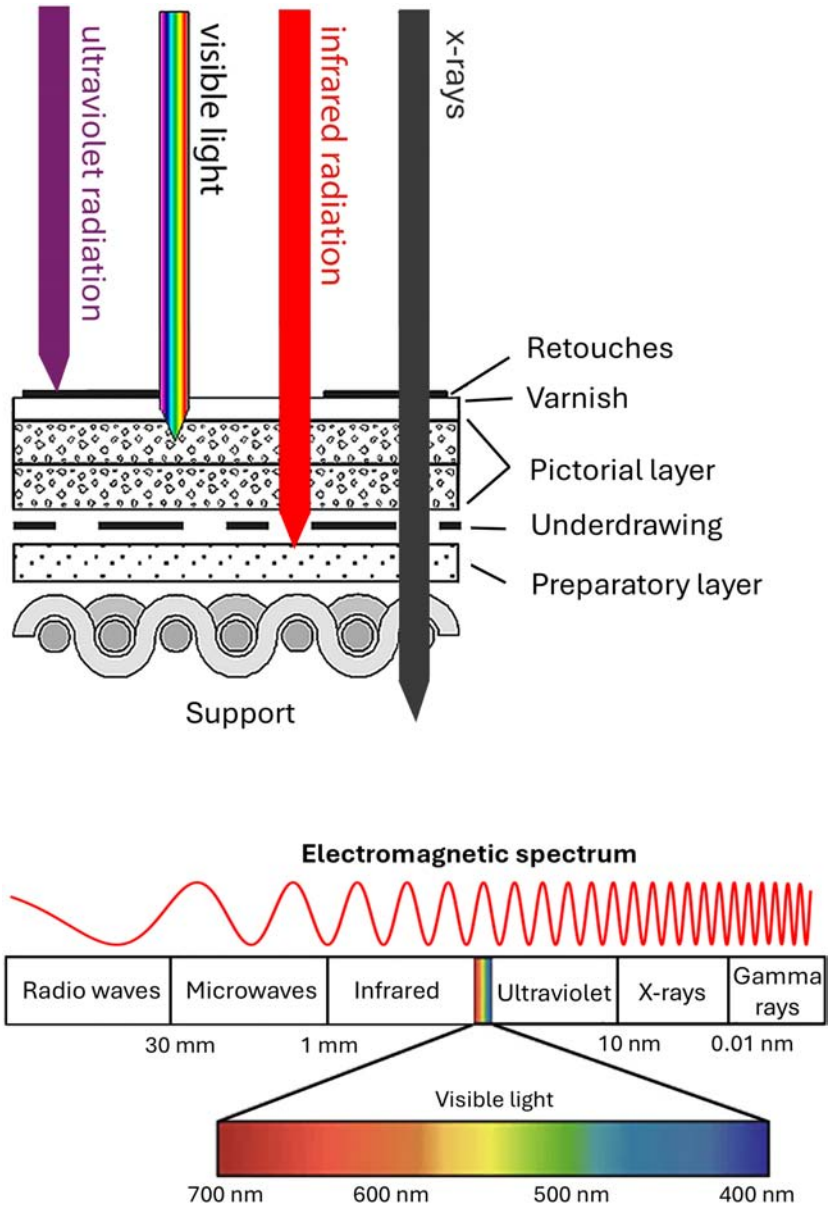

Figure 2.1: Schematic cross-section of a painting illustrating its different layers and the penetration depth of electromagnetic radiation at various wavelength ranges (top). The electromagnetic spectrum, ranging from radio waves to gamma rays, with approximate wavelengths (bottom). Adapted from [37].

or egg yolk for tempera), which is then dissolved in a liquid vehicle (such as turpentine or water), only when necessary. Finally, varnishes are applied to the finished work to protect it from dust, dirt, and environmental factors, while also enhancing

the color and sheen of the surface [38].

2.2 Non-invasive analytical techniques for material identification

When electromagnetic radiation interacts with an object, it can undergo reflection, transmission, absorption, or emission, with the specific proportions determined by the characteristics of the radiation and the physical and chemical properties of the material. Non-invasive analytical techniques exploit these interactions to study materials without causing damage.

In this section, the focus will be on the most commonly used analytical techniques for determining the materials present in historical documents and paintings [39, 40]. While there is a wide range of techniques employed for this purpose, the emphasis will be placed on those that are non-destructive (i.e., they do not require sampling or the destruction of the artwork) and non-invasive (i.e., they do not even require physical contact with the sample), and are also portable techniques, meaning they can be transported to the locations of the artworks, thereby minimizing potential risks to the objects. This approach will allow for a better understanding of the information provided by each technique, as well as the limitations that should be considered when assessing the potential of HSI in this field.

Among these techniques are X-Ray Fluorescence (XRF), X-Ray Diffraction (XRD), Fourier Transform Infrared (FTIR) spectroscopy, and Raman spectroscopy. A summary of these methods, including their basic operating principles, the information they provide, their limitations, and selected references of their application in CH studies, is presented in Table 2.1.

Each technique provides different, and often complementary, information, mak-

Table 2.1: Basic operating principles, information provided, limitations of the most commonly used analytical techniques for determining materials in CH, and references to studies where these techniques have been applied.

Technique	Principle	Information obtained	Limitations	References
X-Ray Fluorescence (XRF)	Radiates X-rays onto the sample, causing emission of characteristic X-rays from elements within it.	Elemental composition.	Cannot identify chemical compounds or molecular structure. The surface must be sufficiently flat. Depth-sensitive analysis.	[41–44]
X-Ray Diffraction (XRD)	Analyzes the diffraction pattern of X-rays interacting with the crystalline structure of the material.	Crystalline structure.	Requires crystalline samples. Each measurement point can take between 30 and 45 minutes. Depth-sensitive analysis.	[44–46]
Fourier Transform Infrared (FTIR) Spectroscopy	Measures the absorption or emission of infrared radiation by molecular vibrations.	Functional groups and molecular structures (organic and inorganic compounds).	Can be affected by sample preparation (e.g., moisture). Issues if the surface has specular reflectance properties. Surface-limited analysis. You must select the appropriate infrared region depending on the compound of interest.	[42, 43, 47]
Raman Spectroscopy	Measures the inelastic scattering of light by molecular vibrations.	Functional groups and molecular structures (organic and inorganic compounds).	Can be sensitive to fluorescence. The laser must be carefully adjusted to avoid damaging the artwork. Surface-limited analysis. Lower sensitivity for low-scattering materials.	[41, 46, 48]

ing a multi-analytical approach convenient in most cases when determining the materials used by the artist. In this context, some instruments have been developed that integrate multiple techniques, such as the Duetto (XRD+XRF) [49].

In addition to the limitations outlined in Table 2.1, a drawback common to most of these instruments is that they only provide spectra from a single point or a very small measurement area. Techniques such as macro-XRF [50], macro-Raman [51], and macro-rFTIR [52] aim to address this limitation by conducting numerous point measurements, scanning the surface in a grid pattern to generate elemental or com-

pound maps. However, these methods are time-consuming. For instance, measuring a 10 × 10 cm^2 area with a 1 mm step size takes around 9 hours for macro-Raman [51], 56 hours for macro-rFTIR on a 150 × 145 mm area, and 22 hours for macro-XRF [52]. As a comparison, HSI can measure the same area in just seconds, achieving even higher spatial resolution.

2.3 Reflectance spectroscopy

Another technique that has proven to be highly effective for pigment identification is Reflectance Spectroscopy (RS). It measures the intensity of radiation reflected, either specularly or diffusely, from a small area on a surface across the ultraviolet to near-infrared range (and beyond) as a function of wavelength. The resulting curve is referred to as the reflectance spectrum.

According to the standard EN ISO9488:1999 [53], spectral reflectance, $\rho(\lambda)$, is defined as the ratio of the radiant flux reflected from a surface, $\phi_r(\lambda)$, to the incident radiant flux, $\phi_i(\lambda)$, both measured in watts:

$$\rho(\lambda) = \frac{\phi_r(\lambda)}{\phi_i(\lambda)} = \frac{\phi_{sample,r}(\lambda)}{\phi_{white,r}(\lambda)} \tag{2.3.1}$$

where all quantities are expressed as a function of wavelength (λ). As the ratio of two homogeneous radiometric parameters, $\rho(\lambda)$ is a dimensionless quantity that typically ranges either from 0 to 1, or as a percentage just multiplying by 100.

To obtain the spectral reflectance of an object, a reference standard is used, typically a reference white with similar reflectance across all wavelengths. The reflectance of this standard, $\rho_{white}(\lambda)$, is known, certified, and traceable. By measuring this standard with the same capturing parameters and illumination/sensing geometry, we can determine how the light is impinging in the sample, so we obtain the

second part of Equation 2.3.1.

In addition to the reference standard with known reflectance, a measurement is taken under the same lighting and geometric conditions as the sample, but by closing the instrument's entrance with a black stopper/cap or closing the shutter. This dark measurement accounts for the dark current of the detector and the sensor noise, therefore, it is considered as a fixed bias that must be subtracted from the final measurement.

Taking all this into account, we have the following relation:

$$\rho(\lambda) = \rho_{white}(\lambda)\frac{I_{sample}(\lambda) - I_{dark}(\lambda)}{I_{white}(\lambda) - I_{dark}(\lambda)} \tag{2.3.2}$$

where $\rho_{white}(\lambda)$ is the spectral reflectance of the reference standard (ideally equal to 1 on each wavelength for a perfect reflective diffuser), and I represents the light counts or sensor responses.

Spectral reflectance measurements from 200 nm to 2500 nm can be obtained using spectroradiometers, spectrophotometers, or Fiber Optic Reflectance Spectroscopy (FORS) in a circular or squared pointwise manner, typically covering a few millimeters or centimeters. These instruments have been used in the field of CH since the 1930s [54,55], enabling the analysis of a wide range of materials, including paintings, manuscripts, textiles, glass, metals, and minerals [56]. Their applications include detecting alteration products (e.g., gypsum, calcium oxalates) [57], monitoring restoration work and color changes [58], identifying colored materials such as pigments and colorants [58–61], and characterizing binders [62]. Although FORS is primarily considered a surface analysis technique, the information it provides varies depending on the spectral range examined (as shown in Figure 2.1). The SWIR and NIR ranges reveal information about underlying layers and the support, the UV range is

employed to study varnishes, and the VNIR range is primarily used for the identification of pigments and dyes along with the SWIR.

Focusing on its application to material identification, reflectance spectra can serve as characteristic signatures of materials. Depending on the spectral range studied, different information can be extracted. The visible and near infrared (VNIR) range (400–1000 nm) provides information on color and electronic transitions [63]. These electronic transitions can be classified as follows:

1. Transitions between delocalized molecular orbitals. These occur in materials with conjugated double bonds and electron donor/acceptor pairs, leading to electron delocalization. The energy levels in this range correspond to the visible spectrum, resulting in intense transition bands. Examples of materials include indigo, alizarin, and curcumin.

2. Charge transfer transitions. These involve electron transfer between molecular orbitals located at different sites of the molecule or crystal.

3. Ligand field transition. Occurring in transition metal ions within inorganic pigments or gems (e.g., azurite, malachite, ruby), these transitions are caused by the interaction between metal ion orbitals and the ligand field.

4. Energy band transitions occurring in metals and semiconductors. In these materials, atomic orbitals combine to form energy bands, with the energy gap between valence and conduction bands determining the absorption characteristics.

Extending the spectral range beyond the visible into the NIR and SWIR regions (up to 2500 nm) allows for the detection of vibrational overtones and combination bands associated with stretching and bending vibrations. These spectral features

are characteristic of specific functional groups, enabling the identification of various compounds [61].

However, applying this theoretical framework to the practical identification of materials in CH presents several challenges. Pigment identification using VIS spectral reflectance is significantly influenced by multiple factors, including (i) the pigment particle size, (ii) the pigment-to-binder ratio, (iii) the type of binding medium, (iv) scattering and self-absorption phenomena, (v) the presence of varnish and dirt on the surface, and (vi) the conservation state of the piece. These factors can contribute to a change in spectral shape, a shift in the bands, and a variation of relative band intensities [64]. Moreover, in the 400-1700 nm spectral range, absorption bands are weaker and more complex than those in the SWIR region, so the direct identification from the reflectance spectra is challenging and the application of pre- and post-processing techniques, as well as more complex algorithms, is highly suitable to achieve a higher level of confidence in the classification of materials.

On the other hand, pure pigments are seldom present in real artworks. When different materials are mixed (that is, two different pigments, two inks, or even pigment or ink with the support), the resulting spectral reflectance reflects the interaction of these components. Identifying the individual materials in complex mixtures is challenging due to overlapping absorption bands and multiple absorptions from the same functional group [61]. To determine each individual component and its contribution to the final reflectance spectrum, unmixing techniques must be applied. This issue will be explored in detail in chapter 6 and chapter 7 of this PhD thesis.

2.4 Hyperspectral imaging

2.4.1 Definition

As explained in the previous section, RS provides the reflectance spectrum of a small area, typically from a rounded or squared area of few squared millimeters or centimeters. However, if we scan an entire surface point by point, it is possible to generate an image where each pixel contains the reflectance spectrum of the sample at this point. This approach, known as Reflectance Imaging Spectroscopy (RIS) or Hyperspectral Imaging (HSI), extends the capabilities of RS by incorporating spatial information.

HSI was originally developed in the late 1970s and early 1980s for Earth observation in the field of remote sensing. This technique combines spectroscopy and spatial imaging to provide images at different wavelengths, from the UV (between 330 and 380 nm) through the VNIR (from 380 nm and up to 1000 nm) to the SWIR, (usually between 900 and 2500 nm), and even extending into the Middle Wavelength Infrared (MWIR) (up to 5000 nm). The result is a hyperspectral cube, or hypercube, containing three-dimensional data: two spatial coordinates and a spectrum for each pixel of the image [65].

HSI is a non-invasive, contactless technique that enables relatively fast data acquisition. The raw data captured by an HSI camera can be processed to obtain spectral radiance, which can then be converted into spectral reflectance or transmittance for each pixel (further details are provided in subsection 2.4.3). This analysis can be performed over an entire object, covering areas from several square centimeters to a few square meters, which offers a key advantage over the methods discussed in section 2.2 and section 2.3. The combination of spectral and spatial information not

only facilitates material identification but also enables detailed mapping.

To obtain this hyperspectral cube, four different acquisition modalities can be applied [66]:

1. Point scanning (whiskbroom). It involves obtaining a spectrum from a single point on the object. This process is repeated for each point, with the object moved in the x and y directions using a computer-controlled 2D scanning stage (see Figure 2.2 (a)).

2. Line scanning (pushbroom). It acquires spectral measurements from a single line of the object at a time. These measurements are simultaneously recorded by an array detector, and either the object or the device is moved line-by-line (see Figure 2.2 (b)).

3. Area scanning (band sequential). It captures an image of the entire object using a two-dimensional detector array for each spectral band (also known as spectral scanning), without needing to move the object. This process is repeated across several spectral bands, resulting in a series of two-dimensional images that form the hypercube (see Figure 2.2 (c)).

4. Single shot (snapshot). It is a rapid acquisition method that captures the entire hyperspectral image in one exposure. It uses a two-dimensional array detector to simultaneously capture the entire object in multiple spectral bands (see Figure 2.2 (d)).

Regardless of the method used for capture, i.e., the acquisition modality employed to obtain the hyperspectral cube, there are different imaging modes [69]. In the reflectance mode, the object is illuminated by a broadband, spectrally-continuous light source covering the VIS and infrared (IR) spectrum. The light reflected by the

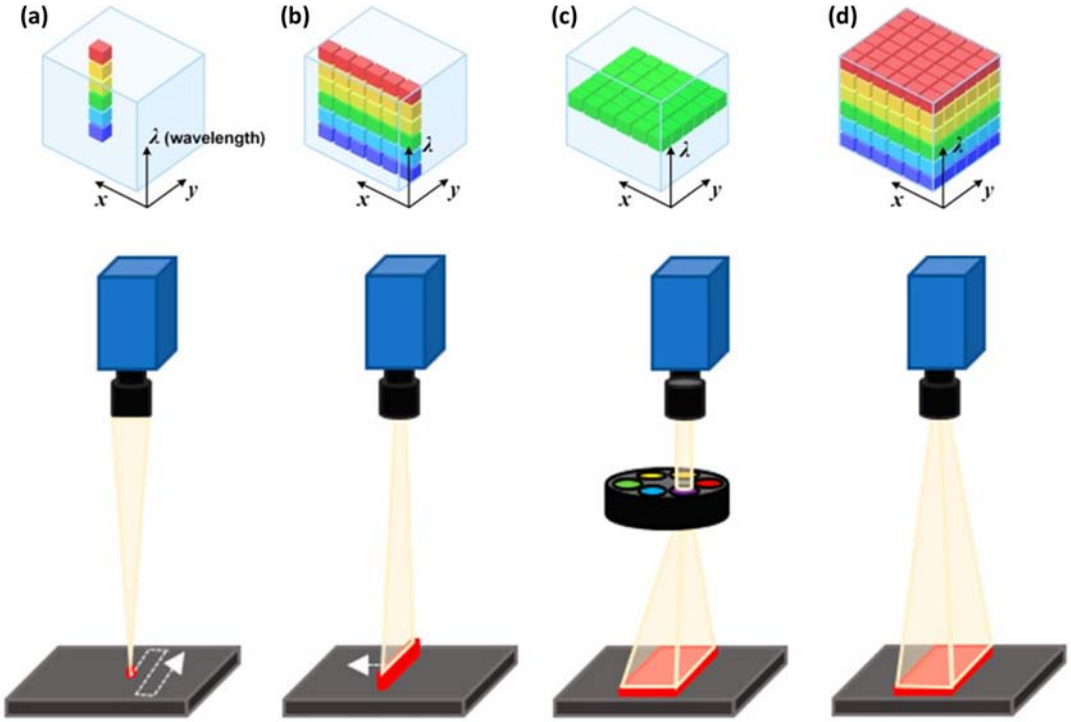

Figure 2.2: Acquisition modalities in HSI: (a) point scanning, (b) line scanning, (c) area scanning, and (d) single shot. Adapted from [67] and [68].

object is captured by the camera, with the wavelengths of the incident light matching those of the reflected light. In transmittance imaging, the object is positioned between the light source and the camera, with only the light that has passed through the object being recorded. In the raking or side-lighting mode, reflectance is captured, but the light strikes the object at a low oblique angle, highlighting surface textures, topography, and indented writing. Lastly, in the luminescence mode, light emission occurs from a substance without heating, as in fluorescence or phosphorescence. The capture process is similar to reflectance imaging, but here, the camera records light that has been absorbed and re-emitted by the object. Different cases arise depending on the illumination and emission wavelengths: ultraviolet-induced luminescence (with UV illumination and emitted radiation in the UV, VIS, or IR range), visible-induced IR luminescence (with VIS illumination and emitted IR ra-

diation), and VIS-induced VIS luminescence (with both illumination and emitted radiation in the VIS range) [70].

2.4.2 Hyperspectral vs. Multispectral vs. RGB imaging

A simpler way to understand the functioning of HSI is to compare it with conventional RGB cameras that we use daily, such as those in our mobile phones, webcams, or digital cameras. RGB cameras capture information in three spectral channels (usually called color channels): red (R), green (G), and blue (B) (see Figure 2.3 (a)). Each of these channels has a relatively wide sensitivity in the electromagnetic spectrum, meaning each channel integrates information from a wide range of wavelengths rather than a specific single wavelength. These regions of the spectrum are known as "spectral bands". The combination of these three bands enables the creation of a color image, mimicking the way the human visual system perceives color.

Imaging spectroscopy, also referred to as spectral imaging, is a general term that includes both multispectral imaging (MSI) and HSI, although the distinction between the two can be ambiguous depending on the literature studied [71]. The primary differentiating factors between these techniques are the number of spectral bands, their bandwidths, and their spectral overlap. MSI systems are designed to capture images in a limited number of spectral bands (i.e. 4 to 20), typically using filters with broader bandwidths ranging from tens to hundreds of nanometers (see Figure 2.3 (b)). In contrast, HSI systems acquire images in a greater number of narrower, contiguous spectral bands, often hundreds or even thousands, with bandwidths typically ranging from 1 to 10 nanometers (see Figure 2.3 (c)). The main advantage of HSI systems lies in their ability to provide nearly continuous spectral measurements, which enhances their accuracy in spectroscopic analysis compared to MSI systems [72]. Generally, HSI systems are designed so that their response can

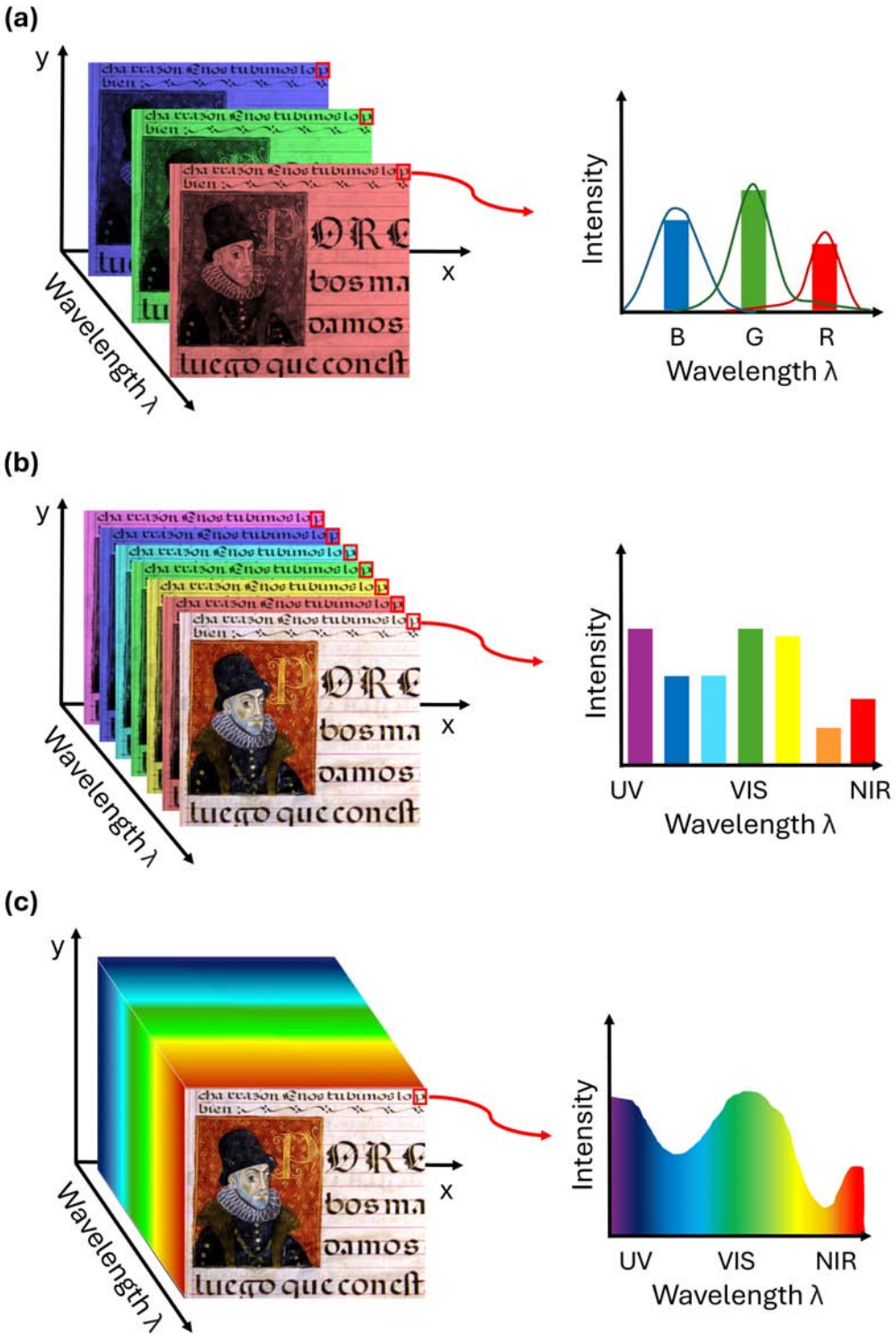

Figure 2.3: Schematic representation of the information provided by (a) RGB, (b) multispectral, (c) hyperspectral imaging.

be considered as a direct spectral measurement, while MSI systems require the application of spectral reconstruction techniques to derive spectra from sensor responses.

2.4.3 Calibration

Similar to the process described for non-imaging measurements in section 2.3, calibration is required to obtain the spectral reflectance or transmittance of an object pixelwise from the camera's response. In this process, raw data is first captured, influenced by three factors: the illuminant used, defined by its spectral power distribution (SPD) which represents the energy emitted at each wavelength for every point in the scene (x, y), $L_e(x, y, \lambda)$; the object to be captured, defined by its spectral reflectance at each point in the scene $\rho(x, y, \lambda)$; and the HSI camera, defined by its sensor sensitivity $S_c(\lambda)$. Therefore, a sensor response is obtained for each wavelength or set of wavelengths within each spectral band $\Delta\lambda$ and each pixel (x, y), which can be defined as:

$$I_{raw}(x, y, \lambda) = \int_{\lambda-\frac{\Delta\lambda}{2}}^{\lambda+\frac{\Delta\lambda}{2}} L_e(x, y, \lambda)\rho(x, y, \lambda)S_c(\lambda)d\lambda \tag{2.4.1}$$

To eliminate the influence of the sensor's sensitivity, $S_c(\lambda)$, radiometric calibration is necessary [73]. This calibration is typically performed by the manufacturer, which provides a radiance mode for converting raw images into radiance images directly.

However, in most cases, the goal is to also eliminate the influence of the illuminant and obtain the spectral reflectance, which depends only on the object being studied, $\rho(x, y, \lambda)$. This can be achieved by following exactly the same equation 2.3.2 if reference images for calibration, dark and white, have been captured beforehand under the same setup (lens, device output bit depth, gain, down-sampling, light

conditions, etc.). In that case, raw captures of the object under study, $I_{sample}(x, y, \lambda)$, can be transformed into reflectance cubes, $I_{ref}(x, y, \lambda)$, using the raw captures of the reference white, $I_{white}(x, y, \lambda)$, and the dark image, $I_{dark}(x, y, \lambda)$, according to the following equation:

$$I_{ref}(x, y, \lambda) = \rho_{\text{white}}(\lambda) \cdot \frac{I_{sample}(x, y, \lambda) \ - I_{dark}(x, y, \lambda)}{I_{white}(x, y, \lambda) - I_{dark}(x, y, \lambda)}, \tag{2.4.2}$$

where $\rho_{\text{white}}(\lambda)$ is the spectral reflectance of the reference tile used for calibration. When the reference tile is also captured covering the entire scene, or the scanned line in pushbroom systems where the sample is moved, flat-field correction can be applied to account for variations in illumination across the field of view, as well as for discrepancies in light sensitivity between the pixels of the camera and differences in the transmission of light through the lens [64].

The illumination/detection geometry for most HSI cameras in reflectance mode is $45^\circ/0^\circ$ for line illumination: light impinges on the sample at 45° and the back-scattered radiation is collected at 0° (all angles are measured with respect to the axis orthogonal to the surface). In practice, however, a floodlight setup is often employed. This method involves using a broad-beamed, uniform light source to illuminate a larger area compared to line illumination, minimizing shadows by incorporating different lamps.

2.5 Hyperspectral imaging for cultural heritage

In the early 1980s and 1990s, wide-band near-infrared reflectography became a well-established technique for the diagnostics of paintings and the identification of details underlying the pictorial layer. For several years, it remained the most effective method for gaining in-depth knowledge about underdrawings, retouchings,

and *pentimenti* (that is, images, forms, or strokes that have been changed and painted over) [64]. Once MSI cameras became a well-established technique in the remote sensing field, their application in CH became possible. An MSI camera was first developed and applied in situ to image paintings at the National Gallery in London, collecting high quality images in a few, broad spectral bands [74]. Initially used for qualitative comparison between bands and to improve image color accuracy, it began to be used for pigment identification in the late 1990s [75, 76]. It was not until the early 2000s that the first application in a museum appeared [77, 78]. The unique, fragile nature of artworks, coupled with the need for portable equipment to reach museums, churches, and conservation centers led to the redesign of the instruments both to safeguard the objects during measurements and to ensure portability.

Since then, MSI and HSI have been used for the examination of hundreds of artworks, including a wide range of famous paintings and documents, such as da Vinci's Mona Lisa [79], Van Gogh's Self-portrait [80], Vermeer's Girl with a Pearl Earring [81], the first draft of the U.S. Declaration of Independence [82], Picasso's Harlequin Musician [83], various works by Goya [17] and by El Greco [84], among others.

Over the years, HSI has gained prominence in this field and has been used for different purposes in TCH [11, 85, 86]. A visual overview of these applications is provided in Figure 2.4. Detailed information on specific applications, including the type of spectral imaging used (MSI or HSI), the spectral range, and the imaging mode (reflectance, transmittance, or raking), can be found in Table 2.2.

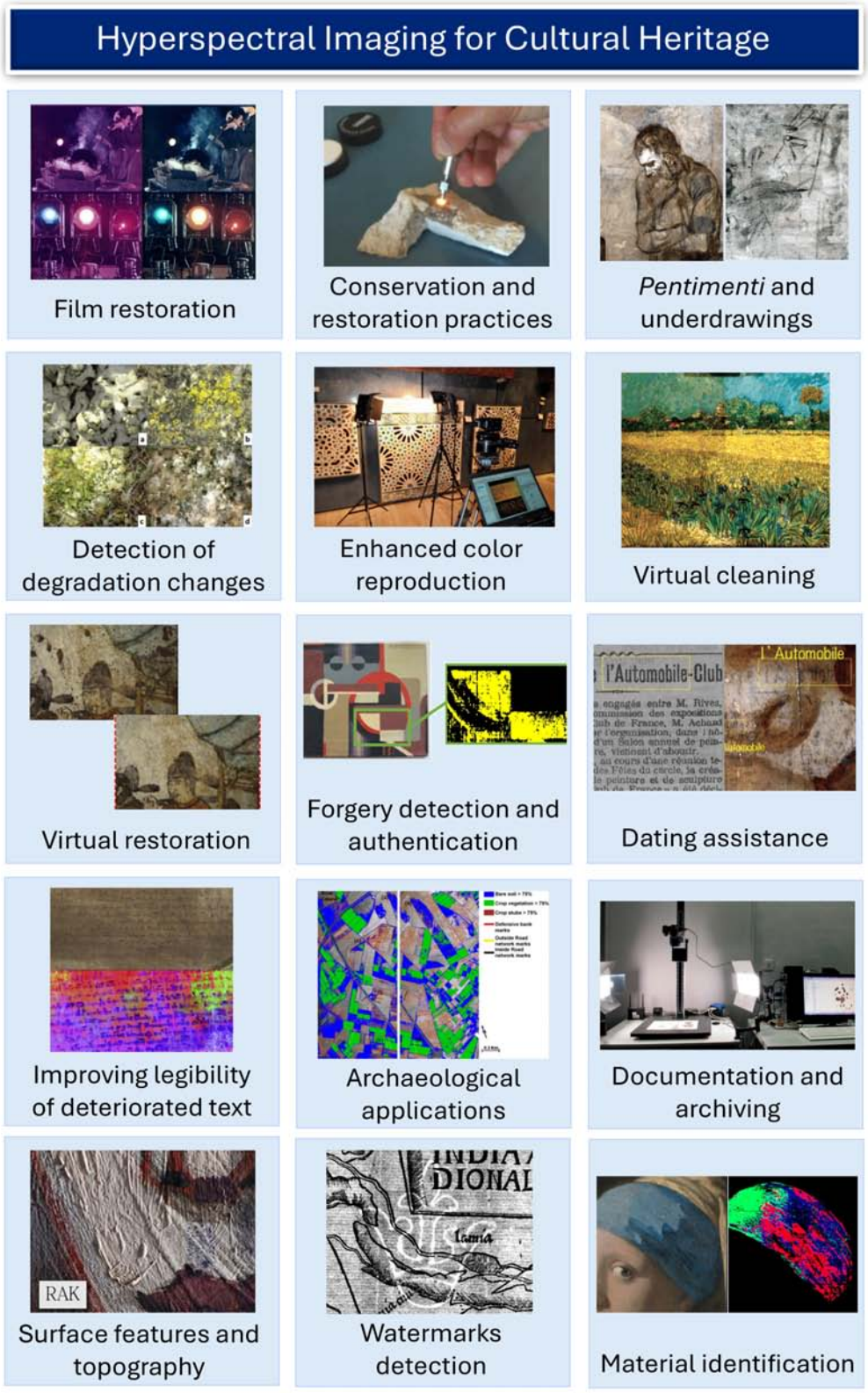

Figure 2.4: Overview of different applications of HSI in CH. Adapted from [16, 81, 83, 87–98].

Table 2.2: Applications of HSI in CH, detailing the field of study, type of spectral imaging used, spectral range, imaging mode, and relevant references.

Field of application	Field of study	Spectral imaging	Spectral range	Imaging mode	Ref.
Conservation and restoration of photographic materials	Restoration of negative and positive films, reconstruction of the supposed original color appearance, and characterization of photographic archival materials.	MSI	400-700 nm	Transmittance	[99]
			380-700 nm		[87]
		HSI	400-900 nm		[100]
Detection and monitoring of restoration and conservation interventions	Evaluation of different consolidants for medieval plasterworks and methods for cleaning of re-polychromed plasterwork. Assessment and control of the laser cleaning process of a historic limestone entrance gate. Evaluation of laser cleaning on paper and parchment. Evaluation of pigment combinations for inpainting or retouching to address issues of metamerism.	HSI	400-1000 nm	Reflectance	[101–103]
			400-1700 nm		[89]
		MSI	320-1550 nm		[104]
			380-730 nm		[105]
Revealing *pentimenti*, underdrawings or sketches, and underlying text	*Pentimenti* and underdrawings revealed in *The Drunkenness of Noah* by Andrea Sacchi, *The Imposition of the Name of S. John the Baptist* by Guido di Piero, the Archimedes palimpsest, *The Tragedy* by Picasso, and *Woman reading* by Stefan Luchian.	MSI	780-1000 nm	Reflectance	[106]
			UV-NIR		[107]
		HSI	950-2500 nm		[9]
			1000-2500 nm		[108]
			360-1150 nm		[109]
			417-1650 nm		[83]

Table 2.2: (continued) Applications of HSI in CH, detailing the field of study, type of spectral imaging used, spectral range, imaging mode, and relevant references.

Field of application	Field of study	Spectral imaging	Spectral range	Imaging mode	Ref.
Detection of degradation-induced changes	Colorimetric analysis for accurate color reproduction and monitoring chromatic variations due to aging. Evaluation of varnish degradation. Monitoring of biological changes on the surface of historical stone.	MSI	400-700 nm	Reflectance	[110]
		HSI	400-1000 nm		[90,91]
Enhanced color reproduction of artworks	Optimal lighting in museums for a natural appearance of artworks in exhibitions and reintegrated pieces.	HSI	380-780 nm	Reflectance	[111,112]
Virtual cleaning	Digital removal (virtual cleaning) of aged or discolored varnish from paintings.	HSI	400-950 nm	Reflectance	[96,113]
Virtual restoration	Inpainting of areas in ancient tomb murals using deep learning.	HSI	380-780 nm	Reflectance	[97]
Forgery detection and authentication	Detection of repaintings, retouches and features related to technical execution. Identification of anachronistic paint, confirming the falsification of two artworks. Identification of fake modern Chinese paintings.	HSI and MSI	950-2500 nm, UV-VIS-NIR	Reflectance	[9]
		HSI	900-1700 nm		[16]
			400-900 nm		[114]
Dating assistance	Hidden newspaper revealed in *Mother and Child by the Sea* by Pablo Picasso, dating the painting to January 1902.	HSI	967-1680 nm	Reflectance	[98]

Table 2.2: (continued) Applications of HSI in CH, detailing the field of study, type of spectral imaging used, spectral range, imaging mode, and relevant references.

Field of application	Field of study	Spectral imaging	Spectral range	Imaging mode	Ref.
Improving legibility of deteriorated text	Enhancement of low-contrast text due to ink bleed, ink corrosion, mold, humidity, water, and foxing. Application to medieval manuscripts. Improvement of carbonized textual writings from scroll collections (such as the Dead sea scrolls and scrolls of Petra and Herculaneum).	HSI	365-1100 nm	Reflectance	[115]
		MSI	360-1000 nm		[116]
			400-720 nm		[117]
		HSI and MSI	400-1000 nm, 350-1150 nm		[95]
Archaeological applications	Remote sensing for archaeological prospection, used to detect surface and subsurface anomalies associated with archaeological structures.	MSI	VIS, NIR, SWIR, TIR	Reflectance	[94,118]
Documentation and archiving	Digitizing paintings for color-accurate image archives. Digitization of the Dead sea scrolls. Important MSI in the case of damaged and barely legible historical documents.	MSI	400-1000 nm	Reflectance	[119]
			445-925 nm		[93]
			650-1000 nm		[120]
Highlighting surface features and topography	Highlighting surface topography in scrolls and palimpsest.	MSI	470 and 910 nm	Raking	[88,93,121]
Detection of watermarks	Imaging watermarks and manuscript content within the support.	MSI	365-870 nm	Transmittance	[88]

2.5.1 Material identification

One of the most widely applied uses of HSI in CH, which is directly related to the main objective of this PhD thesis, is material identification. Early approaches to material identification using MSI and HSI [75,76] involved mock-ups (physical replicas of paintings created using materials and techniques that closely resemble those of the original artwork, allowing for safe experimentation without risking damage to the original) and a sixteenth-century panel painting. In these studies, the VNIR spectral region was used for the segmentation and classification of painted areas, incorporating multivariate analysis techniques such as Principal Component Analysis (PCA).

Multivariate analysis has been widely used in CH applications since the early adoption of HSI, as it facilitates data dimensionality reduction, feature extraction, and image enhancement. PCA, in particular, has proven effective in distinguishing materials in artworks composed of a limited number of components [122]. Additionally, several pigments can be differentiated based on variations in their spectral properties, including spectral slope and the presence or absence of absorption bands [122], without requiring further processing. However, PCA alone is not always sufficient for accurate material identification.

As a result, other techniques such as spectral matching methods, and more recently, machine learning (ML) approaches, have been employed for spectral mapping. In classification tasks, these methods are used to determine the presence or absence of a specific pigment based on a given spectrum. However, when materials are mixed, classification approaches may face significant challenges. In such cases, spectral unmixing techniques can be applied to determine the individual components within a mixture. Unlike classification, spectral unmixing is a regression task,

as it quantitatively decomposes a given spectrum into its base constituents according to a pre-defined mixture model.

2.5.1.1 Spectral matching methods

In spectral mapping, each pixel of a hypercube is classified as a specific material. Each possible material, or class, is represented by a characteristic reflectance spectrum, also known as a spectral signature, which defines the material under study. The collection of these spectra is referred to as the reference or spectral library.

For material identification in artworks, a comprehensive spectral library is typically required. This library should encompass a wide variety of materials, including mixtures in different proportions, as well as different supports and application techniques relevant to the historical period of the artwork. Such a library can be constructed in two ways: by creating mock-ups with well-characterized materials or by acquiring spectra directly from artworks, specifically from areas that have been previously analyzed and whose composition has been confirmed through complementary techniques (as discussed in section 2.2). Additionally, several researchers have developed spectral databases to facilitate this process [123, 124], which are further detailed in subsubsection 2.5.1.5.

Once a spectral library is available, different algorithms can be applied to compare the spectra in a hypercube with those in the reference library. In spectral library search, each spectrum is compared to the library entries, a similarity score is assigned, and the pixel is classified according to the most similar reference spectrum. The comparison algorithm, also known as spectral matching method or spectral metric, quantifies the degree of similarity between the unknown and reference spectra. These methods vary in their sensitivity to shape and magnitude differences between spectra and are broadly categorized into deterministic approaches (based on geo-

metric and physical properties) and stochastic approaches (based on statistical distributions) [125]. Several spectral matching algorithms have been widely used for material identification in CH. A schematic representation of material identification using spectral matching methods is presented in Figure 2.5 (a).

Spectral metrics offer a computationally simple and efficient approach to material classification, making them particularly attractive for CH applications. However, their effectiveness depends on the sensitivity of the selected metric to spectral variations, especially when dealing with aged or mixed pigments [126]. Consequently, a variety of spectral metrics have been employed in previous studies, including Spectral Angle Mapper (SAM) [26, 83, 84, 126–129, 129–143], Spectral Correlation Mapper (SCM) [84, 126, 129, 142, 144], and Spectral Information Divergence (SID) [84, 126, 129, 136]. Other approaches include the Spectral Similarity Scale (SSS) [126], combined metrics such as $SIDSAM$ [126], $JMSAM$ (Jeffries–Matusita distance function combined with SAM) [126], and other combinations [126, 145, 146], as well as Euclidean Distance (ED) [84, 129], Spectral Gradient Mapper (SGM) [84], Binary Encoding (BE) [129], Root Mean Square Error ($RMSE$) [140, 147–149], and the Goodness-of-Fit Coefficient (GFC) [149]. These metrics have been extensively applied in pigment and dye identification and mapping in paintings, illuminated manuscripts, and textiles such as tapestries. The equations for computing some of these spectral metrics are provided in chapter 4, subsection 4.2.3.

However, spectral matching methods become computationally inefficient as the spectral library size increases. While many spectral matching algorithms are effective at identifying spectral shape differences, they may struggle with spectra that differ primarily in magnitude or that are similar in shape for a specific spectral range but correspond to different materials [126]. This challenge is particularly relevant for art historians and conservators, as distinguishing such subtle differences can be crit-

ical for selecting appropriate conservation methods. In such cases, especially when reference spectra include numerous variations for each class (e.g., pigments or inks with different binders, writing supports, or recipes), ML techniques can be used to enhance classification accuracy.

2.5.1.2 Machine Learning algorithms

ML is a field of study in artificial intelligence that focuses on developing algorithms that enable computer systems to learn from data, detect patterns, and make predictions or decisions without being explicitly programmed. ML algorithms are designed to improve their performance over time by analyzing data, identifying patterns, and adapting their models accordingly. In the context of HSI and CH, ML algorithms can learn the spectral characteristics of each material, allowing for more robust classification by accounting for spectral variability across different classes.

ML techniques are generally divided into two categories: supervised and unsupervised learning. In supervised learning, the algorithm is trained on labeled data, meaning it learns from input-output pairs (e.g., classifying pictorial styles from images of paintings). The user provides reference or training data for each class and the algorithm performs the classification task based on this information. In contrast, unsupervised learning involves algorithms that find patterns and relationships within unlabeled data (e.g., clustering or anomaly detection). Unlike supervised classification, unsupervised classification groups classes based on the inherent patterns within the data, without requiring prior labeled examples or training data. A schematic representation illustrating material identification using supervised ML algorithms is presented in Figure 2.5 (b).

To evaluate the performance of supervised ML models, a test set is required. This dataset contains labeled data, similar to the training set, but is never seen by the

model during training. The trained model is applied to this new data, and its predictions are compared against the correct labels, referred to as Ground Truth (GT). Based on these comparisons, a confusion matrix is constructed, which quantifies the model's performance for each class by reporting the number of True Positives (TP; correctly predicted positive cases), False Positives (FP; incorrectly predicted positive cases), True Negatives (TN; correctly predicted negative cases), and False Negatives (FN; incorrectly predicted negative cases). These values are then used to compute key evaluation metrics: precision, recall, accuracy, and F1-score. Precision indicates the reliability of positive predictions (see Equation 2.5.1), recall assesses the model's ability to identify all the positive instances in the dataset (see Equation 2.5.2), accuracy measures the proportion of correct predictions out of all predictions (see Equation 2.5.3), and F1-score is the harmonic mean of precision and recall (see Equation 2.5.4) [150].

$$precision = \frac{TP}{TP + FP} \tag{2.5.1}$$

$$recall = \frac{TP}{TP + FN} \tag{2.5.2}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.5.3}$$

$$F1 - score = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{2.5.4}$$

In a multi-class problem, two approaches can be considered to compute these

performance metrics. The micro-average approach treats all individuals (in our case, the reflectance spectra per pixel) equally, not taking into account differences between the number of instances per class. As a result, micro-average accuracy, precision, recall, and F1-score are identical. The macro-average approach gives each class equal weight in the average, which ensures that performance is balanced across all classes. Macro-average is computed as the arithmetic mean of the metrics for single classes [150]. The micro-average approach weights classes according to their frequency in the dataset, which gives more importance to larger classes. Therefore, poor performance on smaller classes is less impactful as they represent a smaller portion of the overall dataset. In contrast, high macro-average values indicate that the algorithm performs well across all classes, regardless of their frequency. This ensures that each class is considered equally, making it a better measure of performance for imbalanced datasets.

The significant advantage of ML techniques lies in their ability to extract complex, hidden structures from raw spectral data, including both linear and nonlinear relationships. This flexibility has driven the increased use of neural networks (NNs) in pigment identification and classification based on HSI datasets, especially in the last few years. NNs are particularly effective in processing a wide range of data types and extracting patterns that traditional methods may miss, which has led to their successful application in CH research [151]. Their application has extended to the analysis of easel paintings, mural paintings, bronze fragments, illuminated manuscripts, tapestries, tomb murals, and relics.

Some unsupervised models, such as Self-Organizing Maps (SOM) [152, 153], Fuzzy C-means, Density Peak Clustering [154], and k-means clustering [155, 156], have been used in CH to group spectra with similar characteristics, identify anomalies, or as a preprocessing step to select candidate pigments. These models can also

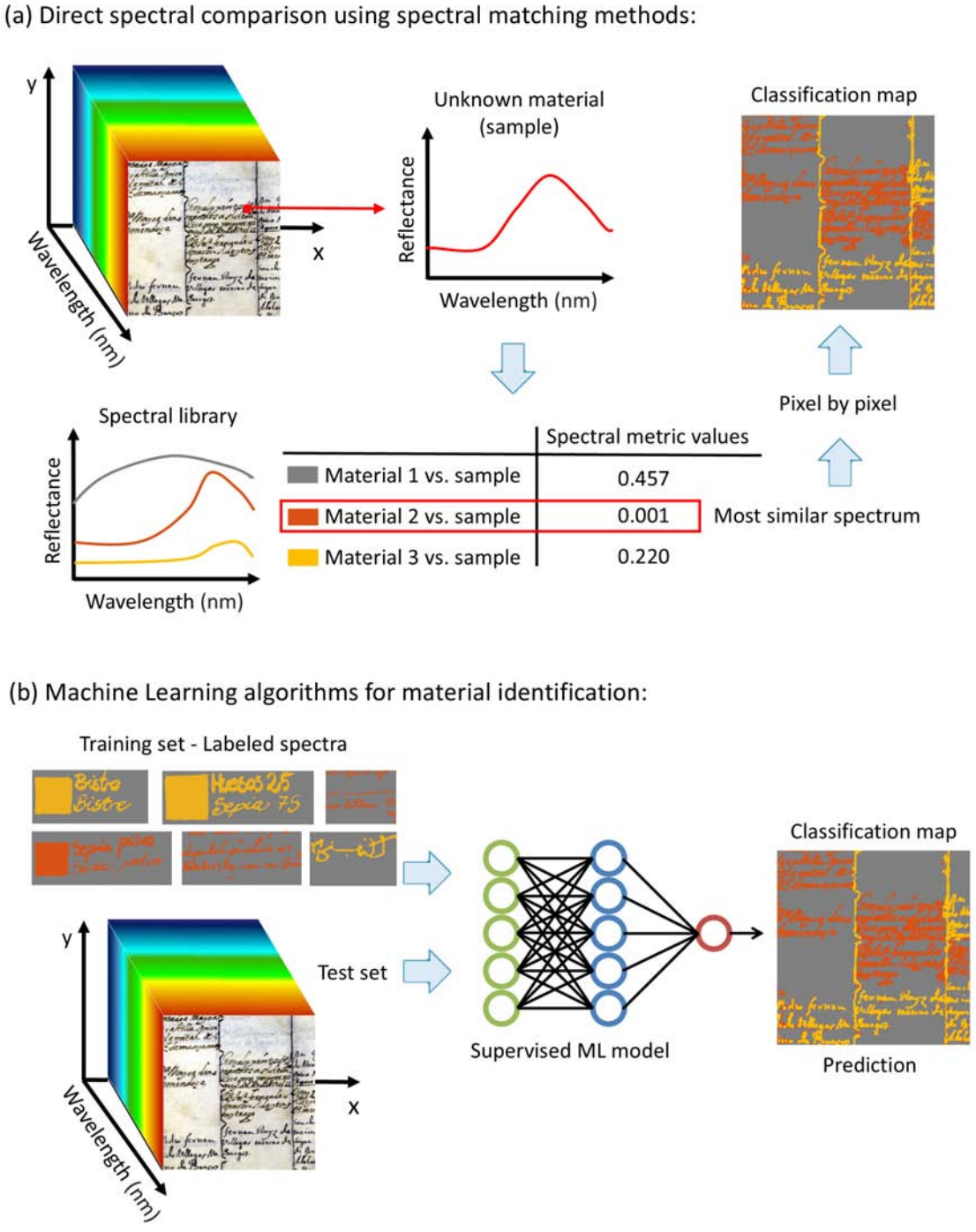

Figure 2.5: Schematic representation of material identification using (a) direct spectral comparison with spectral matching methods, and (b) machine learning algorithms.

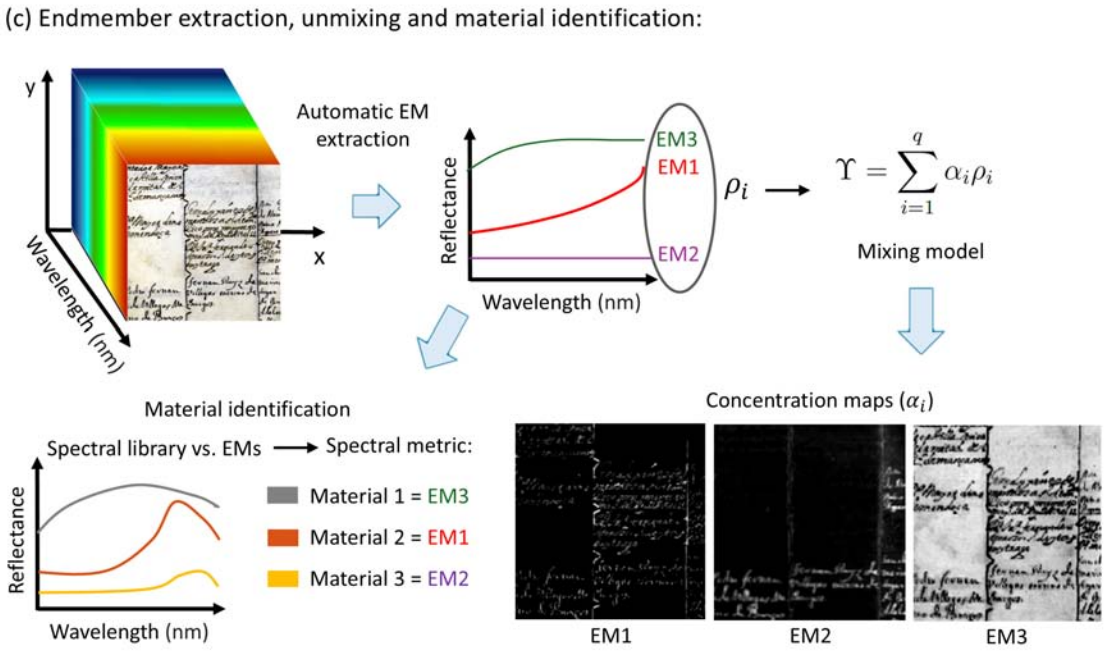

Figure 2.5: (continued) (c) Schematic representation of material identification through endmember extraction and unmixing.

be used for image segmentation, dividing the image into regions with similar spectral characteristics.

For pigment classification, supervised ML algorithms such as Support Vector Machines (SVM) [16, 26, 114, 126, 147, 157–159], Partial Least Squares Discriminant Analysis (PLS-DA) [43, 135, 159], Decision Trees (DT) [147, 159], and k-nearest neighbors (KNN) [159] have been successfully applied. In addition, deep learning (DL) techniques, such as feedforward artificial NN, Fully Connected NN, and 1D-Convolutional NN (CNN) [126], have shown promising results when large datasets are available [160]. For example, Pouyet *et al.* [137] used a Deep NN for pigment identification and mapping using HSI in the SWIR region, and found that the model outperformed traditional spectral matching methods like SAM. Other DL architectures, such as the Multilayer Perceptron (MLP) [26, 152, 161], and CNNs [126, 158, 159, 162] have also been successfully applied in this domain.

However, the main limitation of ML techniques is the large amount of data required to train robust models. These algorithms demand high-quality labeled datasets for accurate classification, and insufficient, noisy, or biased data can severely affect performance. Additionally, issues such as class imbalance, where certain classes are underrepresented, may hinder the model's ability to generalize correctly. Overfitting is another challenge, as models may memorize training data rather than learning generalizable patterns, leading to poor performance on unseen data. Conversely, underfitting occurs when simpler models fail to capture the complexity of the data, resulting in low classification accuracy. Other limitations include the high computational cost and memory requirements of some ML models, as well as their "black-box" nature, where decisions made by DL models may lack interpretability and transparency.

2.5.1.3 Unmixing

Artists have historically experimented with diverse materials to create colors and textures, resulting in a wide range of pigment mixtures with distinct spectral reflectance curves. This makes finding pure pigments in artworks particularly challenging, especially when they undergo aging [163–165], weathering [166], or restoration processes [167]. Additionally, binders and varnishes can modify a pigment's spectral signature [17], while aged varnishes and patinas (thin surface layers formed over time through aging, oxidation, or environmental exposure) further obscure the spectral characteristics of pigments, making direct comparison with spectral libraries insufficient. In many situations, it is not practical to continue extending the number of pigments in the reference library until one covers all the possible mixtures that a given artist (often unknown) might have used. In such cases, advanced analytical methods, such as spectral unmixing and endmember extraction techniques, become

necessary [136, 168].

Spectral unmixing methods involve decomposing a mixed spectrum into its constituent spectra, known as endmembers (EMs), and determining their relative concentrations according to a mixing model $\Upsilon(\lambda) = f(E(\lambda, q), C(q))$ on a per-pixel basis [163], where $E(\lambda, q)$ denotes the spectral library with spectral reflectances for each EM, ρ_i, and $C(q)$ represents the concentration vector of individual abundances, α_i. EMs represent the pure materials used to produce the mixtures, while the concentrations represent the proportion of each EM present in every pixel of the image. The process consists of two main steps: EM extraction and concentration estimation.

EM extraction methods aim to estimate the main spectral signatures from a scene. The Pixel Purity Index (PPI) method seeks to geometrically find the EMs as the vertices of the smallest shape that contains the observed data in an N-dimensional space [163]. NFINDR, on the other hand, deploys the complex of N-dimensional simplex, but the EMs are found iteratively growing the simplex from within the data. The stopping criterion is set so that the simplex connecting the purest pixels is larger than any simplex connecting the pixels found as combinations of others [169]. Nonnegative matrix factorization (NMF) decomposes a data matrix into two lower-rank nonnegative matrices: one representing basis components (endmembers) and the other their corresponding weights (abundances) [170]. Recently, a generative Deep-Learning based model (DeepGun) was introduced for unsupervised unmixing [171] using low-dimensional representations of EMs in the latent space of the generative model. The network is re-trained for each scene and provides a set of EM for each pixel. This model performs better than the parametric non-linear extensions of the linear model, and it is not computationally expensive if subsampling techniques reduce the number of pixels in the spectral image. In addition to automatic EM extraction methods, another possibility is to directly extract the library of EMs

from the painting or object under study by extracting representative areas. Although these areas may consist of mixed pigments rather than pure ones, the resulting spectra can be highly accurate, as they are closer to real spectra than those sometimes obtained through automatic EM extraction methods.

Once the EMs are extracted, the relative contribution of each one to the observed mixed spectra is calculated. This is done using spectral mixing models, which describe how the spectral signatures of different EMs combine to form the observed spectrum. These models describe the physical processes that occur when different pigments are mixed. They can be classified as linear or non-linear. In linear models (see Equation 2.5.5), the mixed spectrum is obtained by a linear combination of EMs weighted by the concentrations:

$$\Upsilon = \sum_{i=1}^{q} \alpha_i \rho_i \tag{2.5.5}$$

where Υ is the reconstructed reflectance of the mixture, q is the number of candidate EMs, ρ_i is the spectral reflectance of the ith EM, and α_i its concentration. This model is commonly used in remote sensing [172–174] and it is considered an acceptable approximation in many real scenarios. Its advantages are physical interpretability, computational tractability, and ease of implementation [175,176]. It has also been used in the field of CH with promising results [168,177].

Nevertheless, when pigments are mixed, the individual components are not discernible with imaging technologies, and scattering effects influence the final spectrum [161, 178]. Therefore, these mixtures are better characterized by a non-linear model [179, 180]. For instance, the Kubelka-Munk model describes the relationship between the absorption and scattering coefficients of incident light in highly light-scattering materials, requiring information about the optical properties of the

materials studied [181]. It has been used in the CH domain as a proof of concept [161, 182], but has not been extended due to memory requirements and computation times [183]. In remote sensing, non-linear unmixing is performed by parametric extensions of the linear model addressing the spectral variability found in the set of EM across different pixels [184] or considering different combinations within the set of EM [185]. These approaches do not incorporate the physical principles that underlie the Kubelka-Munk model. Many studies have focused on developing robust, accurate, and tractable unmixing algorithms [163]. Grillini *et al.* [177] explored various mixing models, including the additive, subtractive, Yule-Nielsen, additive-subtractive, subtractive-additive, LIP (Logarithmic Image Processing) additive and LIP subtractive models, finding that the subtractive model outperformed the others.

After selecting the mixture model and performing unmixing to obtain concentration maps for each EM, the next step is material identification. This is typically done by matching each EM to a material from a reference library using different spectral metrics. Ideally, the reference library should include a broad range of common pigments applied to an appropriate ground layer, as some pigments become transparent in the near-infrared range and the support material also influences the spectrum [134]. The pigments should be selected based on traditional recipes and materials relevant to the period of the artwork. A schematic representation of material identification using EM extraction and unmixing is presented in Figure 2.5 (c).

Several approaches have been proposed for spectral unmixing and pigment identification [26,156,163], including the use of first and second derivatives of the spectra [81,138,164], clustering techniques [175], and DL methods [161,161,162,183,186–189]. The ENVI's software spectral hourglass wizard (ENVI, Exelis VIS, Boulder, CO) has also been used [127, 128, 138, 190, 191], but it is slow and not fully automatic [134].

Although spectral unmixing is a highly valuable technique, it remains an active

area of research, as achieving effective and accurate unmixing continues to present significant challenges. One of these challenges is the lack of correlation between the spectra of the original components in the mixture and the endmembers (EMs) extracted automatically, which can subsequently compromise the identification of pigments based on these EMs.

2.5.1.4 Multimodal approaches

Although spectral imaging has been used independently for identifying pigments and dyes, a multimodal approach is often preferred by many researchers, as it offers a rapid and definitive way to identify ancient pigments. While HSI alone is a robust technique, some authors claim that it is not exhaustive for the complete characterization of composite painted systems in artworks [64]. To address this limitation, HSI has been effectively combined with other spectroscopic techniques to measure a wider range of properties. Such combinations with other analytical techniques (as discussed in section 2.2) include Raman spectroscopy [17, 132, 192–197], XRF [17, 83, 192–199], FTIR [83, 128, 195], and FORS [128, 196–200], especially for obtaining spatially resolved material information or conducting preliminary material identification, enabling large-area analysis.

As outlined in the objectives of this PhD thesis, the focus is on developing HSI processing techniques to automatically identify the materials used in historical documents and paintings, without relying on other techniques. While other measurement devices have been employed to validate our approach, they are not intended to complement it.

2.5.1.5 Hyperspectral libraries

As previously discussed, identifying materials requires databases, whether for comparison through spectral matching methods with our hypercube, for training ML models, or for pigment identification in the final step of unmixing.

Several databases of spectral reflectance for pigments, dyes, binders and varnishes have been proposed and used in the past, typically derived from measurements of mock-ups using FORS [123, 201, 202]. Regarding MSI or HSI databases, there are a few examples. For instance, A. Cosentino proposed an MSI pigment database using 18 bandpass filters covering the 400-925 nm range [203]. This database, which includes the Pigments Checker STANDARD v.5 of CHSOS, contains 69 pigments applied on cardboard. A more comprehensive HSI pigment database was later developed by Deborah *et al.* [124], featuring 195 pigment patches from Kremer color charts applied on acid-free 180-gram paper, covering the 400 to 1000 nm range.

Despite these advances, to our knowledge, no HSI database exists with the purpose to assist in material identification for historical documents. To further advance the field, the availability of datasets with annotated ground truth is essential. Such datasets are critical for training and evaluating ML and DL models to perform different tasks in historical document analysis [204]. However, the creation of GT data is often labor-intensive and costly. Additionally, datasets must encompass sufficient diversity to enable models to generalize effectively, including documents of varying origins, materials, and historical periods.

CHAPTER 3

HYPERDOC - HYPERspectral database of historical DOCuments and mock-ups

This chapter presents an HSI database of historical documents and mock-ups captured in the 400–1700 nm spectral range. It was developed to address the gap in the field, as no existing database of this nature was found. Specifically, it was designed to facilitate the use of HSI for analyzing historical and artistic documents, with a focus on material identification and mapping. It is the outcome of specific objectives O1 and O2 of this PhD thesis. As shown in Figure 1.3, this was one of the first objectives addressed in the thesis and proved to be one of the most time-consuming due to the extensive efforts required. However, the successful collaboration among a multidisciplinary team made its completion possible.

This study has been submitted to the journal *Scientific Data* as:

A. B. López-Baldomero* *et al*. "HYPERDOC - HYPERspectral database of historical DOCuments and mock-ups from 400 to 1700 nm", 2025.

Additionally, part of this work has been published as a conference paper:

A. B. López-Baldomero*, E. Valero, A. Reichert, F. Moronta-Montero, M. Martínez-Domingo, and A. López-Montes, "Hyperspectral database of synthetic historical inks," in *Archiving Conference*, vol. 21, pp. 11–16, Society for Imaging Science and Technology, 2024.

3.1 Introduction

To preserve historical documents and enhance their accessibility and understanding, more than 60 digital image databases have been developed [205]. These serve as

resources for diverse image processing applications, including layout analysis [206], binarization (i.e., separating ink and support into binary values) [207, 208], content analysis [209,210], author identification [211], and improved readability of degraded documents [212].

Most existing databases comprise digital images acquired using conventional RGB cameras. However, HSI cameras have gained increasing prominence in the field of CH in recent years, as they record images in a relatively fast and non-invasive way, providing far more information [11, 85, 86].

In the context of document analysis, HSI and MSI have demonstrated significant advantages over conventional methods. For instance, binarization using HSI [23] or MSI [213–220] data achieves improved separation of ink and support compared to RGB imaging. In forensic analysis, spectral data have enabled the detection of ink mismatches, aiding in the identification of document alterations or forgeries [221, 222]. HSI has also been used for material identification, such as inks [223, 224] and pigments [26, 72, 85, 126], as further detailed in subsection 2.5.1.

Spectral databases play a crucial role in material identification. Previously proposed MSI and HSI databases are discussed in subsubsection 2.5.1.5. The MSI database by A. Cosentino [203] includes the Pigments Checker STANDARD v.5 from CHSOS, which contains 5 black pigments out of 69 pigments applied on cardboard. From the data, it was shown that browns and blacks lack sufficient spectral features for identification. Therefore, MSI in the 400-925 nm spectral range is not always adequate to perform material recognition tasks. A more comprehensive HSI pigment database by H. Deborah *et al.* [124], which includes 195 pigment patches from Kremer applied on acid-free 180-gram paper, features 14 black pigments.

Despite these advances, existing spectral databases primarily focus on pigment samples applied to modern supports and lack the diversity required for historical

document analysis. They do not account for variations in writing supports, ink formulations used across different regions and periods, or the effects of aging. The inclusion of real historical documents is crucial to ensure the applicability of these databases to real-world scenarios. Additionally, datasets must be sufficiently diverse to enable ML and DL models to generalize effectively, and the availability of annotated ground truth (GT) is essential for their training and evaluation [204]. However, the creation of GT data is often labor-intensive and costly.

As part of this PhD thesis, an HSI database of historical documents and mock-ups captured in the 400-1700 nm spectral range was created. This database was developed within the framework of the HYPERDOC project [225], which focuses on using HSI to analyze historical and artistic documents for material identification and mapping. A distinction must be made here: the complete HYPERDOC database, which is available online on the Color Imaging Lab group's website [225], includes a total of 1681 samples. However, this chapter focuses on describing and analyzing only the samples within this database that meet three specific criteria: (i) they are spatially registered in both the VNIR and SWIR spectral ranges, providing a complete spectrum from 400 to 1700 nm for each pixel; (ii) they include GT data; and (iii) the materials present in the GT are known [226]. In total, this subset consists of 720 samples.

The workflow for data collection, capture, and post-processing is summarized in Figure 3.1. Ink mock-ups were created using historical recipes and materials [227, 228], including metallo-gallate inks, sepia, carbon-based inks, and mixtures, which were applied on five different supports. The database also includes pencil mock-ups and historical inks subjected to artificial aging. Additionally, historical documents from the 15th to 17th centuries were sourced from the Archive of the Royal Chancellery and the Provincial Historical Archive of Granada (Spain). Two

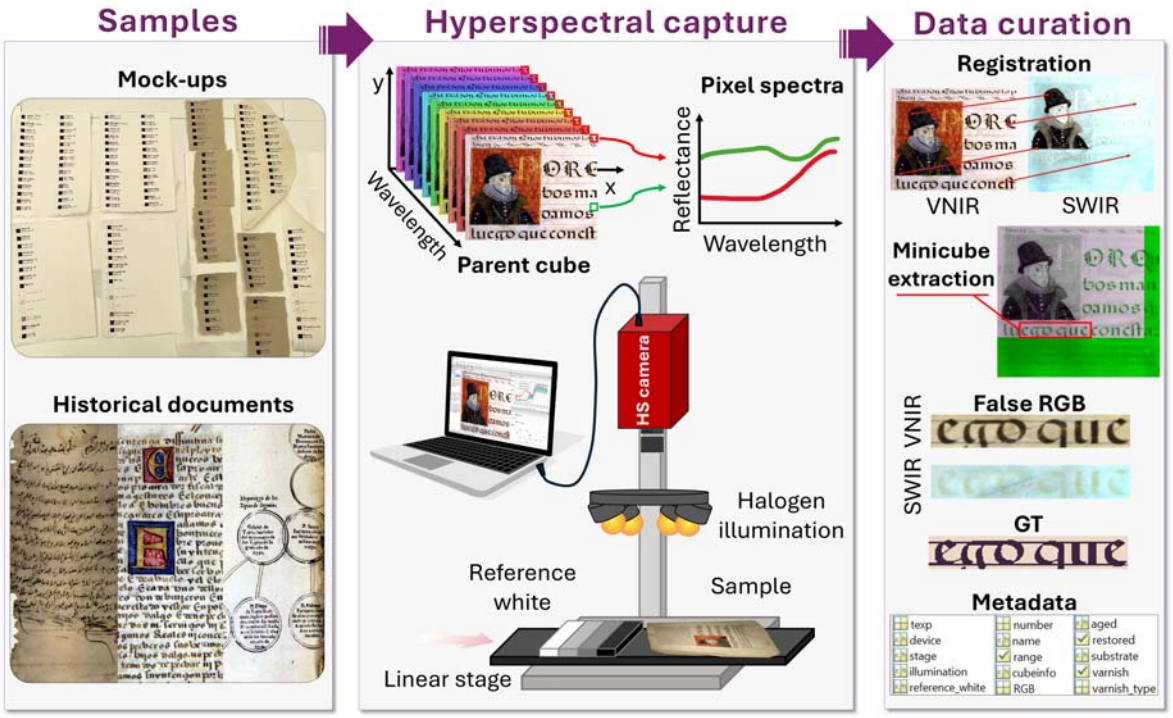

Figure 3.1: Graphical abstract of the capture methodology and data curation.

spectral ranges were captured: 400–1000 nm and 900–1700 nm, which were spatially registered. From these captures, the full hypercubes (referred to as parent cubes) were cropped into smaller, representative Regions of Interest (*minicubes*) to facilitate faster data processing. GT annotations were created to label the materials present at the pixel level. False-color RGB images were generated for both spectral ranges, and Metadata was also integrated into each parent cube and minicube to provide detailed information.

The HYPERDOC database stands as a unique and versatile resource, integrating mock-ups and historical documents to support a wide range of applications in HSI and historical document analysis. Its subsets have been utilized in studies addressing diverse tasks such as ink classification using ML and DL techniques [229, 230] (detailed in chapter 5 and subsection 3.4.1), binarization to enhance text legibil-

ity [231, 232], spectral unmixing to identify the materials [233] (expanded in chapter 7), and colorimetric analysis of aging processes in mock-ups [234, 235]. Thanks to the variety of recipes in mock-ups, they can be used to perform multivariate exploratory analysis in order to identify spectral features related to some kinds of inks or their components, and spectral changes related to the amount of ink deposited. It also facilitates comparative studies between mock-ups and historical documents, including analyses of artificial versus real aging, the state of conservation of the real documents in comparison to other samples, and the impact of different supports or writing instruments on spectral properties. The spectral data also allow simulations of document appearances under various illuminants.

Moreover, the HYPERDOC database fosters interdisciplinary collaboration between the image processing and restoration-conservation communities, encouraging the adoption of advanced techniques such as HSI, which remains underutilized in practical applications within archives and museums. This database holds significant potential to drive innovation in the restoration and preservation of our CH.

3.2 Methods

3.2.1 Samples description

3.2.1.1 Mock-ups

Mock-ups of historical inks on different supports

39 mock-up samples of inks applied on five different supports were prepared following historical recipes from the 13th to the 17th centuries and bound with Arabic gum [236], resulting in a total of 195 samples. The contour of a 1 x 1 cm square was drawn with pencil and filled with ink, and two lines of text were written with brush

(top) and fountain pen (bottom) (see Figure 3.2 (a)). This allows for the study of spectral variations based on the amount of ink deposited and enables the assessment of different writing instruments. The inks used include 8 recipes of metallo-gallate inks, 4 of sepia, 6 carbon-based inks, and 17 mixtures. In addition, 4 samples of constituent materials used in some of the ink recipes were made. The exact quantities of ingredients used to prepare the inks are detailed in the column *'general info'* of the file *'HYPERDOC_database_info.ods'* included in the database.

The inks used to create the mock-ups are listed below:

- Metallo-gallate inks (or iron gall inks). Ancient recipes for iron gall ink were followed [236, 237]. In the preparation of this kind of ink, first, 23.53 g of oak apples (from Kremer Pigmente GmbH) were crushed (but not pulverized), wrapped in a cotton cloth, and macerated in 400 ml of water under direct sunlight for 3 days. After that, the solution was filtered, obtaining the gallic acid. Primarily, pulverized ferrous sulfate ($FeSO_4$-$7H_2O$) was used and added to gallic acid. We varied the ratio of gallic acid and ferrous sulfate obtaining 3 different inks. In addition, pulverized copper sulfate ($CuSO_4$) and zinc sulfate ($ZnSO_4$) were mixed separately with ferrous sulfate to create different inks. Furthermore, Andalusian ink recipes [237] were followed including in both pomegranate juice, but with the addition of myrtle leaves in the gallic acid only in one of them. In addition, the pigment Atramentum (Kremer Pigmente GmbH) was used. It is produced from the reaction of tannic acid extracted from oak bark and iron salts, and has historically been referred to as 'ink stone'.

- Sepia inks. Two types of sepia ink were used in this database: one extracted directly from the animal and another obtained in powder form from Kremer Pigmente GmbH. For the natural extraction, three samples were prepared: pure

ink diluted with water, and two others varying the concentration of Arabic gum binder.

- Carbon-based inks. The database includes several carbon-based inks from Kremer Pigmente GmbH, obtained by charring different materials:
  - Ivory black: it was originally obtained from charred ivory or horns, although nowadays it is based on bone black, being the best quality of that ink. This is the most intense, deep black pigment.
  - Bone black: produced by the carbonization of bones, usually with temperatures over 400°C but not above 800°C.
  - Lamp black: obtained by collecting the condensed smoke produced from a flame fueled by mineral oil, tar, pitch, or resin within brick chambers.
  - Grape seed black: for an extended period, different types of fruit stones have been charred and used as pigments. This pigment is obtained by carbonizing grape seeds, and it has a bluish tone.
  - Cherry black: it is similar to grape seed black, but in this case cherry pits are charred, obtaining a warmer, brownish black.
  - Bistre: it is a warm, deep transparent brown pigment, traditionally obtained from the soot of beech wood or other trees.
- Mixed inks. The database includes 85 samples (17 inks on 5 supports) of different mixed inks. Sepia ink (the one obtained from the animal) was mixed with pure iron gall ink, lamp black, and bones black, in three different proportions: 25:75, 50:50, and 75:25. In addition to sepia, iron gall ink was mixed with lamp black and bones black separately in the same three proportions: 25:75,

50:50, and 75:25. Andalusian red earth (from Kremer Pigmente GmbH) was also mixed with iron gall ink and lamp black in a 50:50 ratio.

- Constituents. Finally, 4 reference samples were made with constituent materials of different inks, including Arabic gum diluted at 20%, pomegranate juice, myrtle leaves infusion, and Andalusian red earth, bound with an Arabic gum solution.

All inks were applied to five types of support, selected for their historical relevance: three types of handcrafted paper from Paperlan® made of 100% cotton fiber, 100% linen fiber, and a linen/cotton mixture 50/50%, hemp paper from Wanderings®, and goatskin parchment from Forum Traiani®. These supports were selected based on those commonly found in historical documents [238].

Pencil mock-ups

This subset of mock-ups includes 14 pencil types from Faber Castell® with varying grades of hardness (8B, 7B, 6B, 5B, 4B, 3B, 2B, B, HB, F, H, 2H, 4H, and 6H) applied to 4 different supports: cotton-linen, cotton, linen, and hemp paper. Similarly to the ink mock-ups, a 1 x 1 cm square was filled with pencil and a line of text indicating its hardness was written (see Figure 3.2 (b)). In total, 56 samples were created.

Mock-ups of artificially aged metallo-gallate inks

Three variants of metallo-gallate inks were deposited on hemp paper, including pure iron gall ink, iron gall ink with copper sulfate, and a mixture of iron gall ink and lamp black in a 50:50 ratio. These inks were used to create squares, strokes, and drops, and subsequently subjected to artificial aging using two distinct methods (see Figure 3.2 (c)). In the first method, an aging chamber (Solarbox® 3000 eRH, Neurtek)

was used following the norm ISO 5630-3 (1996). The chamber operated at a temperature of 80 C°, a relative humidity of 65%, and a radiation of 550 W/m^2. Samples were extracted and captured after 0, 72, 144, and 288 hours of aging, corresponding to 0, 3, 6 and 12 days, respectively. In the second method, aging under acidic conditions was studied by exposing samples to hydrochloric acid vapors for 72, 144, and 288 hours.

3.2.1.2 Historical documents

Manuscripts of the Provincial Historical Archive of Granada

This subset comprises five different documents preserved in the collection of Arabic documents at the Provincial Historical Archive of Granada [239]. Four of these are notarial documents dating from 1488 to 1494, while the fifth is an undated religious text. All five contain handwritten text (see example in Figure 3.2 (d)). Scanning Electron Microscopy with Energy Dispersive X-ray Spectroscopy (SEM-EDX) has identified various types of inks in these documents, including mixed iron gall ink with earth pigments, mixed carbon-based ink with earth pigments, pure carbon-based inks, and pure iron gall inks. The support used in all the documents has been confirmed to be linen paper, as determined through a combination of optical microscopy, Scanning Electron Microscopy (SEM), and FTIR [239,240].

Illuminated manuscripts from the Archive of the Royal Chancellery of Granada

This collection comprises seven documents on parchment, containing lawsuits of nobility dating from 1459 to 1608 [241] (see example in Figure 3.2 (e)). Different pigments and dyes are present in certain areas of these documents; however, these regions were not included in the main focus of this database, which is primarily on inks. The inks used in the handwritten text were identified as iron gall ink with

different sulfates, specifically zinc (Zn) and copper (Cu). Ink identification was performed using XRF, while the inorganic elements used in the preparation of supports were identified using a combined XRD and XRF system [242].

Family tree book from the Archive of the Royal Chancellery of Granada

This series of eight documents from the 16th and 17th centuries comprises family trees, predominantly handwritten with some stamped sections (see example in Figure 3.2 (f)). All documents have cotton-linen paper as the support. Previous analyses identified two types of ink: carbon-based ink and a mixture of sepia and iron gall ink. The documents were restored in 2005 through mechanical cleaning with non-greasy soft rubbers, washing in water, and drying under weight and blotters. The ink types were characterized using SEM by the conservators in charge.

3.2.2 Hyperspectral imaging capture

Two line-scan HSI cameras from Resonon Ltd. (Bozeman, Montana, USA) were used, together with the associated software Spectronon Pro 3.5.5: the Pika L and the Pika IR+. Details about the spectral range covered by each camera, number of spectral channels, spectral resolution, number of pixels per scanned line, spatial resolution at 60 cm, maximum frame rate, and F-number for each camera are provided in Table 3.1. These cameras operate on a push-broom technique, capturing images line by line, which requires either the movement of the object or the camera to scan the entire scene. For image acquisition, a linear translation stage from Resonon Ltd. was used along with 4 stabilized halogen lamps positioned to minimize specular reflections and placed at 30 cm from the documents. To ensure controlled lighting conditions, all other lights in the room were turned off. A video illustrating the capture process is available at the following link: [243]. The optimal exposure time

was determined using the 90% reflectance patch from the Sphere Optics Zenith Lite Multistep of size 20x20 cm, or a Teflon bar with known reflectance, serving as the reference white. Then, the software automatically adjusted the scanning speed and camera data acquisition to ensure 1:1 vertical:horizontal aspect ratio. To maintain the reference white and the document at the same distance from the camera, magnets and additional supports were used, as shown in Figure 3.2 (g). The distance between the camera and the samples was approximately 60 cm for the VNIR camera and 40 cm for the SWIR camera, resulting in linear fields of view (swath) of 13.5 cm and 14.5 cm, respectively. This setup yielded an estimated spatial resolution of 0.15 mm/pixel for the VNIR range and 0.227 mm/pixel for the SWIR range. Spectral binning was performed during capture to enhance the signal-to-noise ratio, resulting in 150 and 168 bands for the VNIR and SWIR captures, respectively. Before capturing the documents, reference images for calibration were acquired. To convert raw data into reflectance values, dark subtraction and flat-field correction were applied. These steps ensure that variations in illumination and sensor response across the field of view are accounted for, eliminating system-induced artifacts from the data. For this purpose, 30 lines of the reference white, the 90% reflectance patch from the Sphere Optics Zenith Lite Multistep or a Teflon bar, were captured. The mean value of these 30 lines, calculated pixelwise along the longitudinal axis, was then used as the reference white to correct non-uniformities in illumination and determine the light incident on the sample. A dark reference image was also captured by blocking the light entering the camera, allowing the removal of intrinsic sensor noise caused by dark currents. All captures were saved in BIL format including a metadata header file (HDR format). Further details on the conversion of raw captures to reflectance data are provided in next subsection.

Figure 3.2: Examples of samples from different subsets: (a) mock-ups of historical inks on cotton paper, (b) pencil mock-ups, (c) mock-ups of artificially aged metallogallate inks, (d) manuscript of the Provincial Historical Archive of Granada, (e) illuminated manuscript from the Archive of the Royal Chancellery of Granada, (f) family tree from the Archive of the Royal Chancellery of Granada; and (g) HSI capture using the Pika IR+ camera.

Table 3.1: Specifications for the Pika L and Pika IR+ HSI systems.

Parameter	Pika L	Pika IR+
Spectral range (nm)	400 - 1000	900 - 1700
Spectral channels	300	368
Spectral resolution - FWHM (nm)	2.7	5.6
Pixels per scanned line	900	640
Spatial resolution at 60 cm (mm/pixel)	0.15	0.34
Max frame rate (fps)	249	240
f/#	2.4	1.8

3.2.3 Data curation

3.2.3.1 Reflectance from raw

Raw captures of the spectral cubes were transformed into reflectance cubes using Equation 2.4.2 in subsection 2.4.3. This transformation is performed pixelwise in the camera software before cube storage, except for the multiplication by $\rho_{white}(\lambda)$, as the software assumes it to be 100% at every wavelength.

Once the reflectance cubes were in BIL format, they were converted to MAT format using MATLAB (Release R2023a, The MathWorks, Inc., Natick, MA, USA). The code used for the transformation is available on GitHub [244]. It is during this step that the reflectance of the object is multiplied by the reflectance of the reference white ($\rho_{white}(\lambda)$). During this process, linear interpolation was applied to ensure a consistent 5 nm sampling interval across both cameras, resulting in 121 bands between 400 and 1000 nm for the VNIR and 161 bands between 900 and 1700 nm for the SWIR range.

3.2.3.2 Registration

Spatial registration is performed to align pixelwise the captures in the VNIR and SWIR ranges and equalize the spatial resolution. In this case, the SWIR capture was used as the reference image, while the VNIR capture, with its higher spatial resolution, was transformed to minimize artifacts in the final registered image. The registration was performed using one band from the VNIR (700 nm) and one band from the SWIR hypercubes (1000 nm). These bands were selected based on preliminary trials, and their position below 1200 nm in the SWIR range. The latter condition was set to avoid proximity to the onset of the high reflectance region of metallo-gallate inks in the SWIR range, which could lead to a lack of key points necessary for proper

registration. Feature-based image registration with SURF features [245] was used within the MATLAB Registration Estimator App (Release R2023a, The MathWorks, Inc., Natick, MA, USA), applying either an affine or projective spatial transform. The registration quality was assessed using overlay images and the Structural Similarity Index Measure (SSIM) [246], after testing different features and spatial transforms, to ensure satisfactory registration. The final registration transformation was then applied to all spectral bands within the VNIR cube. After the process, the parent cubes are obtained, that is, the hyperspectral images of the full pages from which the minicubes are extracted later on. These parent cubes are included in the folder 'ParentCubes' in the HYPERDOC database [226].

3.2.3.3 Minicube extraction

We define a *minicube* as a crop extracted from a full document or page, with sizes ranging from [34 x 33] to [181 x 508] pixels, selected from representative areas of the full documents. Spectral images delivered by the HSI devices are stored as spectral cubes (i.e. hypercubes), usually of extremely large size of even gigabytes of data per capture. Thus, the extraction of minicubes facilitates faster processing of spectral information, as explained before. To highlight the impact of this approach, the parent cubes from which the minicubes are extracted range in size from [368 × 143] to [2485 × 615] pixels. This corresponds to a reduction of up to 94% in the number of pixels, significantly improving data processing efficiency. Each minicube contains data from one or two inks, the support, and sometimes pencil markings. Minicube extraction was performed on the registered VNIR and SWIR cubes using identical spatial coordinates, with Regions of Interest selected based on areas where different inks or materials were present in the document. These minicubes are included in the folder 'minicubes' in the HYPERDOC database [226].

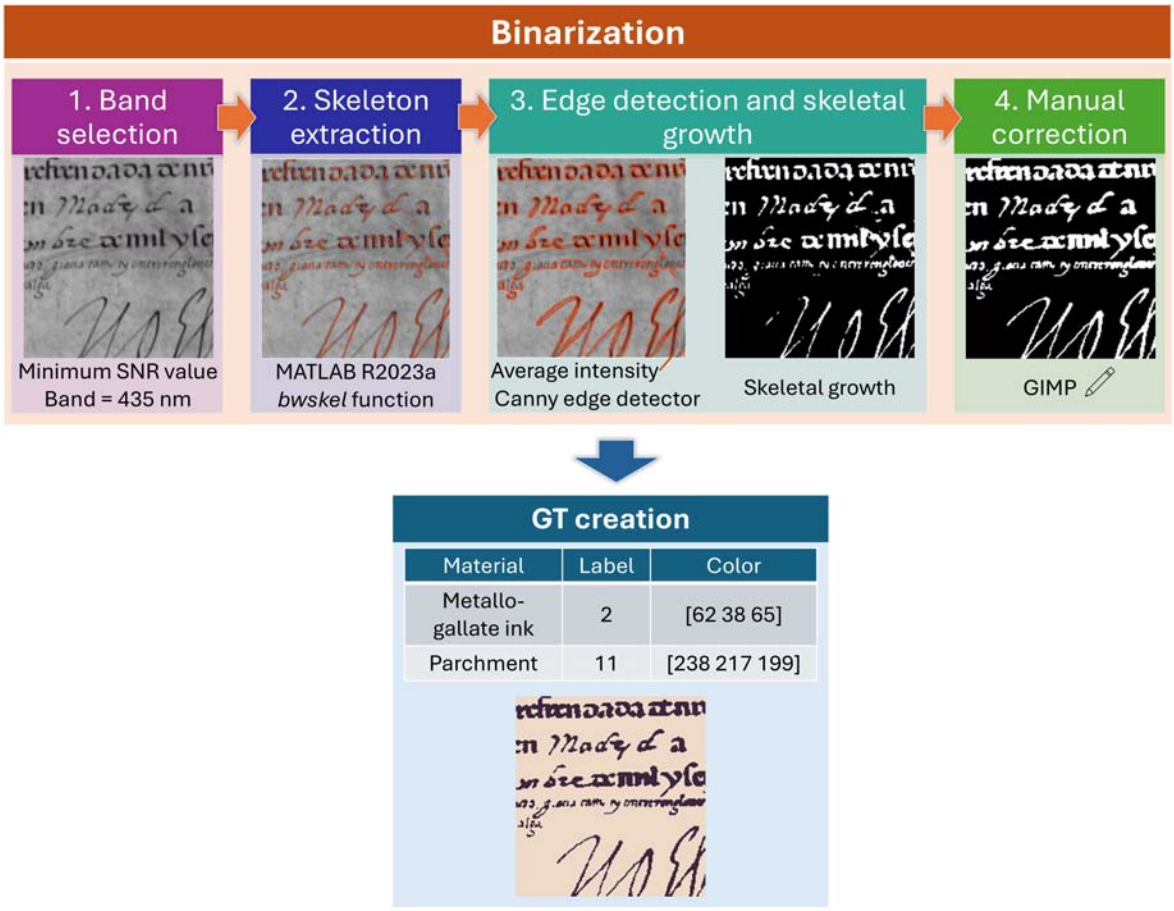

Figure 3.3: Steps involved in the creation of the Ground Truth (GT) images.

3.2.3.4 False RGB images and Ground Truth creation

For each parent cube and minicube, a false-color RGB image was generated by assigning specific spectral bands to the [R, G, B] channels. For the VNIR range, bands [50, 34, 9], corresponding to wavelengths of 645 nm, 565 nm, and 440 nm, respectively, were used, yielding a color appearance similar to that observed in the sample. Similarly, for the SWIR range, bands [141, 61, 21], representing wavelengths of 1600 nm, 1200 nm, and 1000 nm, were selected [226].

A GT image was also created for each minicube using a semi-automatic method. The process is illustrated in Figure 3.3 and consists of two main steps: binarization,

which separates the foreground (pigments and inks) from the background (support), and GT creation, which assigns labels to differentiate between materials. The binarization process involves four steps. First, a band with high contrast between the ink and the background was selected by searching for the minimum Signal-to-Noise Ratio (SNR) value. Second, the skeleton of the part of the image covered with ink was extracted using the MATLAB R2023a function bwskel, based on Lee *et al.*'s medial surface axis thinning algorithm [247]. Third, the skeleton width was adjusted until the intensity of surrounding pixels matched the average intensity of the borders of the ink covered region, extracted by the Canny edge detector. This is a variation of the method proposed by Ntirogiannis *et al.* [248], where the skeleton was manually corrected and then forced to grow until it met those borders. Fourth, manual correction using the open-source software GIMP was performed after obtaining the binarized image, by visually comparing the result with a false RGB image of the minicube. After completing these steps, a binary image was generated with two labels: 0 (background) and 1 (foreground). The GT was then created by assigning different labels to distinguish materials in the binary image. The mapping between labels, GT colors, and materials is provided in the file *'Materials_label_and_colormap_assignation.ods'*. The final indexed images, including the index map and associated colormap, were saved as PNG files in the folder 'GT' [226]. GT images were not created for the parent cubes due to difficulties in providing accurate pixel-level annotations for these larger and more complex regions.

3.2.3.5 Metadata info

Detailed information about each sample is included as metadata within the minicubes. In total, 24 attributes were included, which can be divided into three main categories according to the type of information provided: sample information,

capture information, and other relevant data.

The sample information group contains 11 attributes: identifier number, name, general information about the sample (ink components, origin or recipe), type of support, height in pixels, width in pixels, number of bands, wavelengths captured, date of production of the document, aging status (either naturally or artificially aged, or not aged at all), and restoration status (whether restored or not).

The capture information group contains 6 attributes: device used, range captured, stage, exposure time, type of illumination, and reference white used.

The other relevant information group contains 7 attributes: colormap of the GT, GT labels, parent cube name, and pixel coordinates used to extract the minicube within the parent cube. Using the GT from the PNG files and GT labels, spectra of pixels belonging to each class are averaged and the mean and standard deviation is stored, along with the number of pixels used in the average.

For the parent cubes, metadata were also included. In this case, as GT images are not available, only 16 attributes were included, excluding the identifier number and all attributes in the 'other relevant information' group.

3.3 Data Records

The HYPERDOC database [226], comprising hyperspectral images of historical documents and mock-ups, is publicly available on Google Drive (and will be hosted in the *figshare* repository following the journal's revision process) as part of the Hyperdoc project [225]. The data is structured as shown in Figure 3.4 and includes the following folders and files:

- Folder *minicubes* - Hypercube files with metadata (in HDF5 format): This folder contains HSI datacubes for each minicube, captured in the VNIR and SWIR

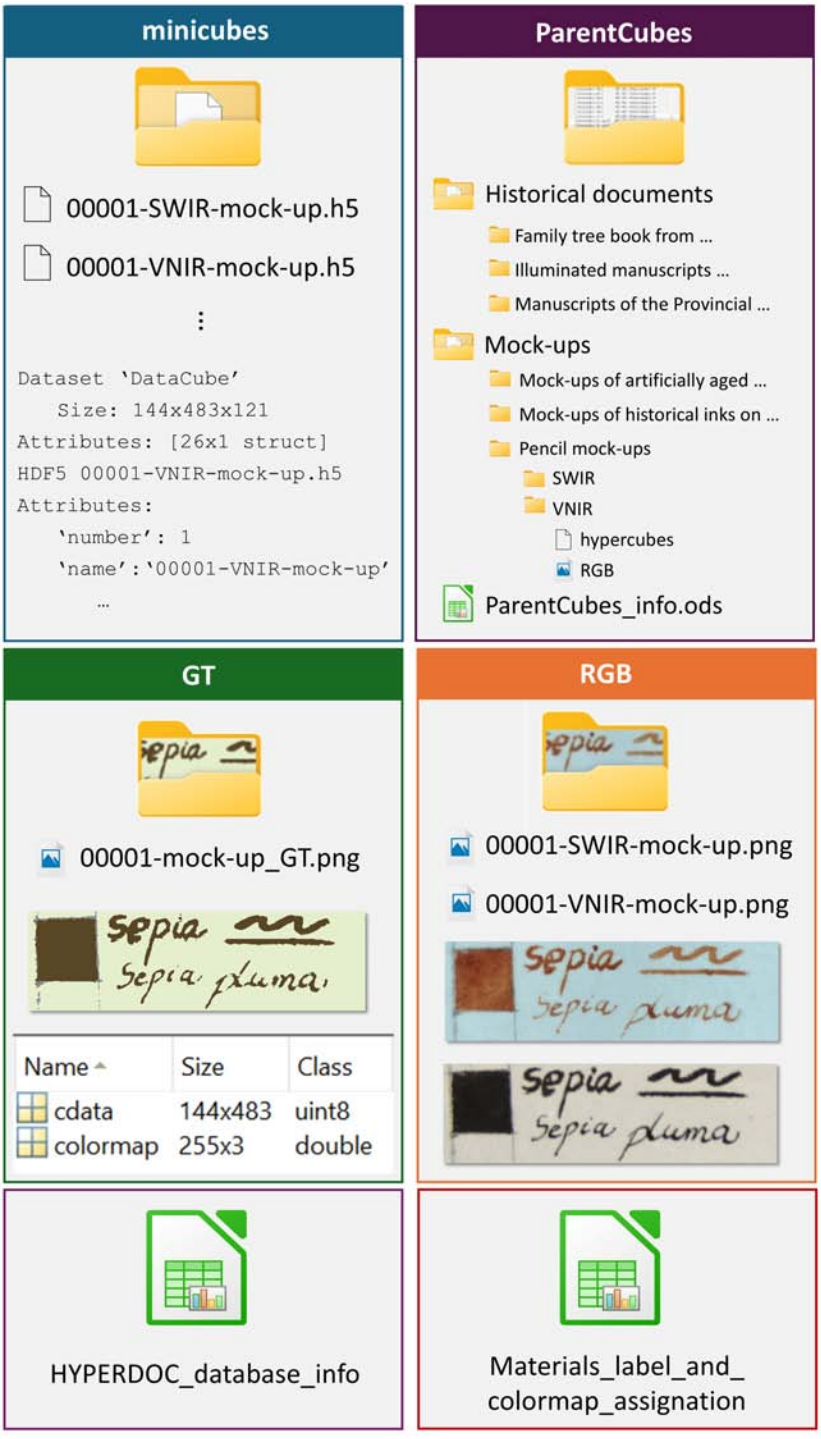

Figure 3.4: Folder and file structure in the database.

spectral ranges, along with associated metadata. The datacubes are stored as a 64-bits double-precision floating-point matrices with dimensions M x N x λ, where M and N correspond to the spatial dimensions of the image, and λ represents the number of spectral bands or wavelengths. Each datacube includes metadata with 24 attributes, describing key information such as acquisition settings and sample details. A detailed description of these attributes, including their data types and possible values, is provided in Table 3.2. Three spectral range options are defined, including ultraviolet and visible (UVIS), although no samples within this range are present in the database. This option has been added to accommodate UVIS samples present in the complete database published on the Color Imaging Lab's website [225].

- Folder *RGB* - False RGB images (in PNG format): This folder contains false-color RGB images for both the VNIR and SWIR ranges, generated using the methods described in subsubsection 3.2.3.4. These images provide a convenient way to quickly visualize the minicube content.

- Folder *GT* - Ground Truth images (in PNG format): This folder includes GT images, where pixel values are directly mapped to colormap indices that relate to RGB data specific for each material and assigned according to a pre-existing list of materials present in the database. Each file contains an indexed image stored in the variable *'cdata'*, along with the associated colormap. The correspondence between material types and their respective indices and RGB values in the colormap is documented in the file *'Materials_label_and_colormap_assignation.ods'*.

- Folder *ParentCubes*: This folder is organized in two subfolders, one for each data set: 'Historical documents' and 'Mock-ups'; and an OpenDocument Spreadsheet (in ODS format) named *'ParentCubes_info.ods'*, which contains par-

ent cubes information for quick reference, extracted from the metadata. Each folder has its own subfolders for each data subset. Within those subfolders, there are the 'VNIR' and 'SWIR' folders, which contain the hypercubes in .h5 format, including the datacube and metadata with 16 attributes, as well as the 'RGB' folder with false RGB images in PNG format.

- File *'Materials_label_and_colormap_assignation.ods'* - OpenDocument Spreadsheet with material labels and colormap assignments (in ODS format): This file contains the mapping between material types (e.g., inks, pencils, or supports) and their corresponding indexed values in the GT images. The colormap assigns RGB values (0–255) to each material index. For each minicube, the associated materials can also be found in the *'GTLabels'* attribute within the metadata.

- File *'HYPERDOC_database_info.ods'* - OpenDocument Spreadsheet with hypercube information (in ODS format): The spreadsheet provides essential information about the minicubes for quick consult extracted from the metadata, which includes the set they belong to (mock-ups or historical documents), subset, name of the minicubes, name of the parent cubes, coordinates within the parent cube used for minicube extraction, general information about the minicube, materials found, support, date, information about if it is aged, restored, or not, and finally, exposure time and reference white used during capture.

Table 3.2: Description and data types of attributes in the Metadata field within each minicube, categorized into three groups: (1) sample information, (2) capture information, and (3) other relevant data.

Group	Metadata	Description	Data type
1	number	Identifying number, from 00001 onwards.	String
1	name	Name of the minicube.	String
1	cubeinfo	Information about the type of sample (mock-up or historical) and relevant details such as materials or recipes used.	String
1	substrate	Type of substrate or support: parchment or paper, and fiber type.	String
1	height	Height of the minicube in pixels.	32-bits unsigned integer
1	width	Width of the minicube in pixels.	32-bits unsigned integer
1	bands	Number of spectral bands in the minicube.	32-bits unsigned integer
1	wl	Wavelengths or spectral bands captured.	32-bits unsigned integer
1	date	Date of creation of the mock-ups or historical documents (year or century, depending on the information available).	String
1	aged	Indicates whether the document is aged. Possible values: 'No' (not aged), 'Art.' (artificial aging, with details of the method and hours), and 'Nat.' (naturally aged).	String
1	restored	Indicates whether the document has been restored. Logical value.	8-bits unsigned integer
2	device	HSI camera used in the capture.	String
2	range	3x1 logical indicating the capture range: 1 0 0 for UVIS; 0 1 0 for VNIR; 0 0 1 for SWIR.	8-bits unsigned integer
2	stage	Translation stage used to perform the capture. All captures were done using the 'linear' stage.	String
2	texp	Exposure time during capture in milliseconds.	64-bits double precision floating point

Table 3.2: (continued) Description and data types of attributes in the Metadata field within each minicube, categorized into three groups: (1) sample information, (2) capture information, and (3) other relevant data.

Group	Metadata	Description	Data type
2	illumination	Illumination used during the capture. Halogen lamps were used in all cases.	String
2	reference_white	Reference white used to calibrate reflectance measurements. Possible values: 'Teflon' (for the teflon bar) and 'Multi_90' (the 90% reflectance patch from the Sphere Optics Zenith Lite Multistep).	String
3	GT_cmap	RGB values from 0 to 1 associated with each index in the GTs. Size 16 x 3.	64-bits double precision floating point
3	GTLabels	Materials associated to indexes used in GT.	String
3	parent_cube	Hypercube from which the minicube was extracted.	String
3	position	Coordinates used to extract the minicube from the parent cube: [xmin xmax ymin ymax].	32-bits unsigned integer
3	spectra_mean	Mean spectra of all pixels associated with the same index in the GT. Size: λ x number of indexes in the GT.	32-bits single precision floating point
3	spectra_std	Standard deviation of the mean spectra. Size: λ x number of indexes in the GT.	32-bits single precision floating point
3	pixels_averaged	Number of pixels associated with the same indexes in the GT and used to calculate the mean. Size: 1 x number of indexes in the GT.	32-bits unsigned integer

Table 3.3 summarizes the number of minicubes and the total pixel counts associated with each material type across the different subsets. For a visual representation, Figure 3.5 presents the distribution of minicubes (top) and the distribution of pixels on a logarithmic scale (bottom) across subsets and classes in the database using horizontal bar graphs. To the right of 0, the minicubes correspond to historical

documents, while to the left, they belong to the mock-ups set. A higher number of minicubes is included in the mock-up category compared to historical documents, representing 73% of the total database. In terms of pixel count, mock-ups account for nearly 94% of the total.

The high number of minicubes containing pencil is due to the first two subsets of mock-ups (historical inks on different supports and pencil samples), where all samples include pencil. However, this material is completely absent in the set of historical documents. Other materials or classes not present in the historical document subsets include cotton, hemp, pure andalusian red earth, pure sepia, mixture of carbon and sepia, and a combination of metallo-gallate ink with carbon-based ink. Among historical documents, the most represented ink class is pure metallo-gallate ink. For the supports, linen, cotton-linen, and parchment are equally distributed in terms of minicube count; however, in pixel count, linen is less represented compared to the other two. In mock-ups, red earth and its mixtures are among the least represented materials, reflecting their limited historical use.

Table 3.3: Classes and total number of minicubes and pixels in each set and subset (commas used as thousands separators). #: number; mc.: minicubes: Arch.: Archive.

Set	Subset	Class	# of mc.	# of pixels
Mock-ups	Historical inks on different supports	Metallo-gallate ink pure	40	612,949
		Metallo-gallate ink mixture carbon	30	452,478
		Metallo-gallate ink mixture sepia	15	235,795
		Metallo-gallate ink mixture earth	5	82,876
		Carbon pure	30	429,217
		Carbon mixture earth	5	64,782
		Carbon mixture sepia	30	423,009
		Sepia pure	20	300,988
		Pencil	180	183,024
		Andalusian red earth pigment	5	89,075
		Parchment	36	1,204,442
		Cotton-linen	36	1,217,969
		Linen	36	1,456,972
		Hemp	36	1,229,443
		Cotton	36	1,433,623
	Pencil	Pencil	56	153,920
		Cotton-linen	14	121,414
		Linen	14	106,705
		Hemp	14	99,962
		Cotton	14	119,084
	Artificially aged metallo-gallate inks	Metallo-gallate ink pure	18	95,255
		Metallo-gallate ink mixture carbon	10	51,403
		Hemp	28	203,886
Historical documents	Manuscripts Provincial Historical Arch.	Metallo-gallate ink pure	20	13,930
		Metallo-gallate ink mixture earth	5	3,501
		Carbon pure	4	1,942
		Carbon mixture earth	3	2,434
		Linen	36	82,664
		Metallo-gallate ink mixture unknown	4	1,896
	Illuminated manuscripts Royal Chancellery Arch.	Metallo-gallate ink pure	29	94,981
		Parchment	29	251,723
	Family tree book Royal Chancellery Arch.	Metallo-gallate ink mixture sepia	24	29,109
		Carbon pure	23	16,654
		Cotton-linen	31	196,182

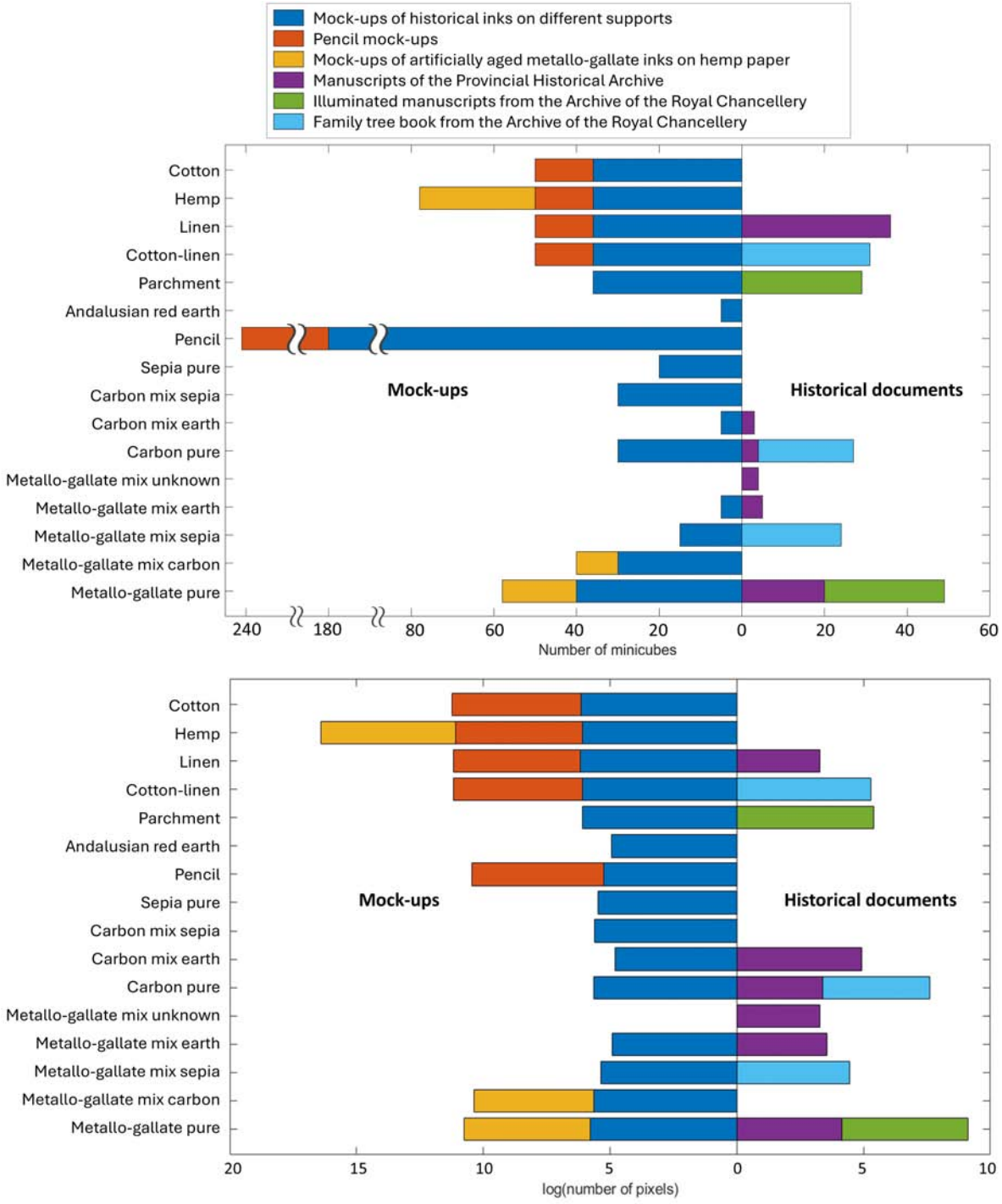

Figure 3.5: Horizontal bar graphs showing the distribution of data: number of minicubes (top) and number of pixels on a logarithmic scale (bottom) by class and subset.

3.4 Technical Validation

Some subsets or samples of this database have been used in previous studies to perform tasks such as classification [229,230], binarization [231,232], spectral unmixing [233], and colorimetric analysis [234,235].

Three categories of inks were classified using ML techniques: pure metallogallate inks, carbon-containing inks, and non-carbon-containing inks [229]. Five traditional ML algorithms—SVM, KNN, Linear Discriminant Analysis (LDA), Random Forest (RF), and PLS-DA—as well as a DL-based model, were trained and evaluated on mock-ups of historical inks applied to different supports and on all the minicubes extracted from historical documents in the present database. All detailed information is provided in chapter 5.

In another study, spectral unmixing techniques have also been applied to identify pure components in mixtures of historical inks, using both mock-ups and selected parent cubes from historical documents included in this database [233]. Mixtures of iron gall, sepia, and carbon-based inks were analyzed by merging VNIR and SWIR ranges. Detailed information can be found in chapter 7.

Binarization tasks were also explored in two studies. In a first study [231], several binarization algorithms including Otsu [249], Niblack [250], Wolf [251], Bradley [252] and a DL based algorithm, were evaluated using a subset of 16th- and 17th-century family trees from the Archive of the Royal Chancellery of Granada. When comparing results from the VNIR and SWIR ranges, the Bradley algorithm consistently produced the best outcomes. A subsequent study [232] compared additional algorithms, including Howe [253], Sauvola [254], and a DL-based algorithm, on mock-ups of inks as well as historical documents not included in this database. While the DL-based approach demonstrated comparable performance to the best traditional

algorithms in both ranges, it provided more consistent and reliable results overall. Comparisons between three-channel images and full HSI cubes containing 121 channels in the VNIR range and 161 channels in the SWIR range revealed that the three-channel images yielded inferior average performance in binarization tasks.

For the subset of artificially aged metallo-gallate inks, colorimetric and spectral analyses were conducted to investigate the aging processes of these inks [234,235].

In addition to the aforementioned applications, a classification task is performed in this chapter using a Bilayered NN to differentiate between four classes in the mock-up set: iron gall ink, non-iron gall ink, support, and pencil (see subsection 3.4.1).

Furthermore, the mean reflectance spectra and standard deviation for each class and subset were calculated to preliminary explore the spectral features in the database (see Figure 3.7 and Figure 3.8 in subsection 3.4.2 for details). Additionally, PCA was used as a dimensionality reduction technique for visualization purposes (see subsection 3.4.3 and Figure 3.9).

3.4.1 Classification of inks and support using a Bilayered Neural Network

A Bilayered NN was trained and tested using the set of mock-ups of historical inks on different supports to classify four categories: iron gall ink, non-iron gall ink, support, and pencil. The iron gall ink category includes both pure and mixed inks containing some amount of iron gall ink. Only VNIR information from samples applied to parchment and cotton-linen paper was used. The Bilayered NN comprised two fully connected layers, excluding the input and the final classification layers. The selection of the model was based on a previous study [126] where a Fully

Table 3.4: Evaluation metrics (accuracy, precision, recall and F1-score) for the four classes: iron gall ink (IGI), non-iron gall ink (NIGI), support (S) and pencil (P).

	IGI	**NIGI**	**S**	**P**
Accuracy	0.67	0.67	0.67	0.67
Precision	0.85	0.65	0.97	0.04
Recall	0.46	0.71	0.70	0.74
F1-score	0.60	0.68	0.82	0.08

Connected NN outperformed a One-dimensional Convolutional NN and an SVM, as well as different spectral metrics, in a pigment classification task. The sizes of the layers were: 32 and 16 respectively, with the hyperbolic tangent as activation function, and softmax activation in the final layer [126]. The model was trained with 85000 spectra of iron gall ink, 85728 spectra of non-iron gall ink, 84000 spectra of support (including parchment and cotton-linen paper), and 21844 spectra of pencil. The ink spectra were extracted from the squared areas of the samples. To test the model, 72 minicubes were used and the performance metrics, accuracy, precision, recall and F1-score, were computed as described in subsubsection 2.5.1.2. To ensure robust performance assessment, a 5-fold cross-validation was conducted. Cross-validation is a technique used to evaluate model generalizability by dividing the dataset into k subsets (or folds). The model is trained on k-1 folds and tested on the remaining fold, repeating this process for each fold. This approach helps mitigate overfitting and provides a more reliable performance estimate. The Classification Learner app included in MATLAB® software (R2023a version) was used to obtain the results for the validation with the pre-trained NN.

As can be seen in Table 3.4, a 67% accuracy (proportion of correct predictions to the total number of input samples) is obtained using the Bilayered NN. The lowest precision is achieved for the pencil class (4%). This is because a lot of support pixels have been misclassified as pencil, as can be seen in Figure 3.6 (A), leading

Figure 3.6: Classification maps of two minicubes: (A) sepia ink, and (B) mixture of iron gall ink and lamp black 50:50, both applied on cotton-linen paper.

to an increased number of false positives for the pencil class. In contrast, the highest precision value (97%) is obtained for the support class, indicating that very few ink or pencil pixels are misidentified as support. Comparing both types of inks, a higher precision but a lower recall is observed for the iron gall ink class. A low recall value indicates a lower proportion of true positives to the actual total number of positive samples. So, an iron gall ink is more likely to be classified as non-iron gall rather than vice versa. When an iron gall ink is mixed with a non-iron gall ink, especially a carbon-based ink, the reflectance of the spectrum diminishes considerably, even with the addition of a minimal quantity of carbon-based ink. This makes the ink spectrum, despite containing a considerable amount of iron gall ink, to become more similar to the spectrum of carbon-based ink. This explains why most of the pixels in Figure 3.6 (B), which shows the classification map for the minicube with an equal mixture of iron gall ink and lamp black (50:50), are classified as non-iron gall ink. Further efforts are needed to improve the overall accuracy, particularly with mixed inks and pencil. This study is part of preliminary classification results, which were later refined and improved, as described in chapter 5.

3.4.2 Average spectra and standard deviation per class and subset

The mean spectra and standard deviation were computed for each foreground material, including pencil, pigments and inks (left plots), and for each support material (right plots), with the results separated by subset (different rows in Figure 3.7 for the mock-ups and Figure 3.8 for the historical documents). These calculations were performed by averaging all the pixels in the minicubes belonging to the same class and subset, using the GT information (example code for performing this calculation in MATLAB and Python is available on GitHub [244]). VNIR and SWIR spectral ranges are presented in the same plot. A noticeable difference in the reflectance spectra is observed in the overlapping region between 900 and 1000 nm, which is a common artifact when data is captured using different sensors. This discrepancy arises from various factors, including differences in spectral bandwidths, low signal-to-noise ratios due to low sensor responsivity in the extremes of the spectra, and slight misalignments in the image acquisition setup, all of which can impact the Bidirectional Reflectance Distribution Function (BRDF) [255].

Ink spectra in the visible range are similar, showing low reflectance values and flat shapes, consistent with the black or brownish appearance of these inks. However, in the near-infrared range, metallo-gallate inks, both pure and mixed with red earth, begin to diverge from other inks, exhibiting a reflectance spectrum increasingly similar to that of the support. This trend is especially prominent beyond 1200 nm, where metallo-gallate inks become nearly transparent. In contrast, carbon-based inks strongly absorb infrared radiation, maintaining low reflectance values. Sepia ink and its mixtures with metallo-gallate allow more infrared transmission but do not reach the near-total transparency observed in metallo-gallate inks. Mixtures of carbon-based inks with other pigments, such as sepia or red earth, significantly re-

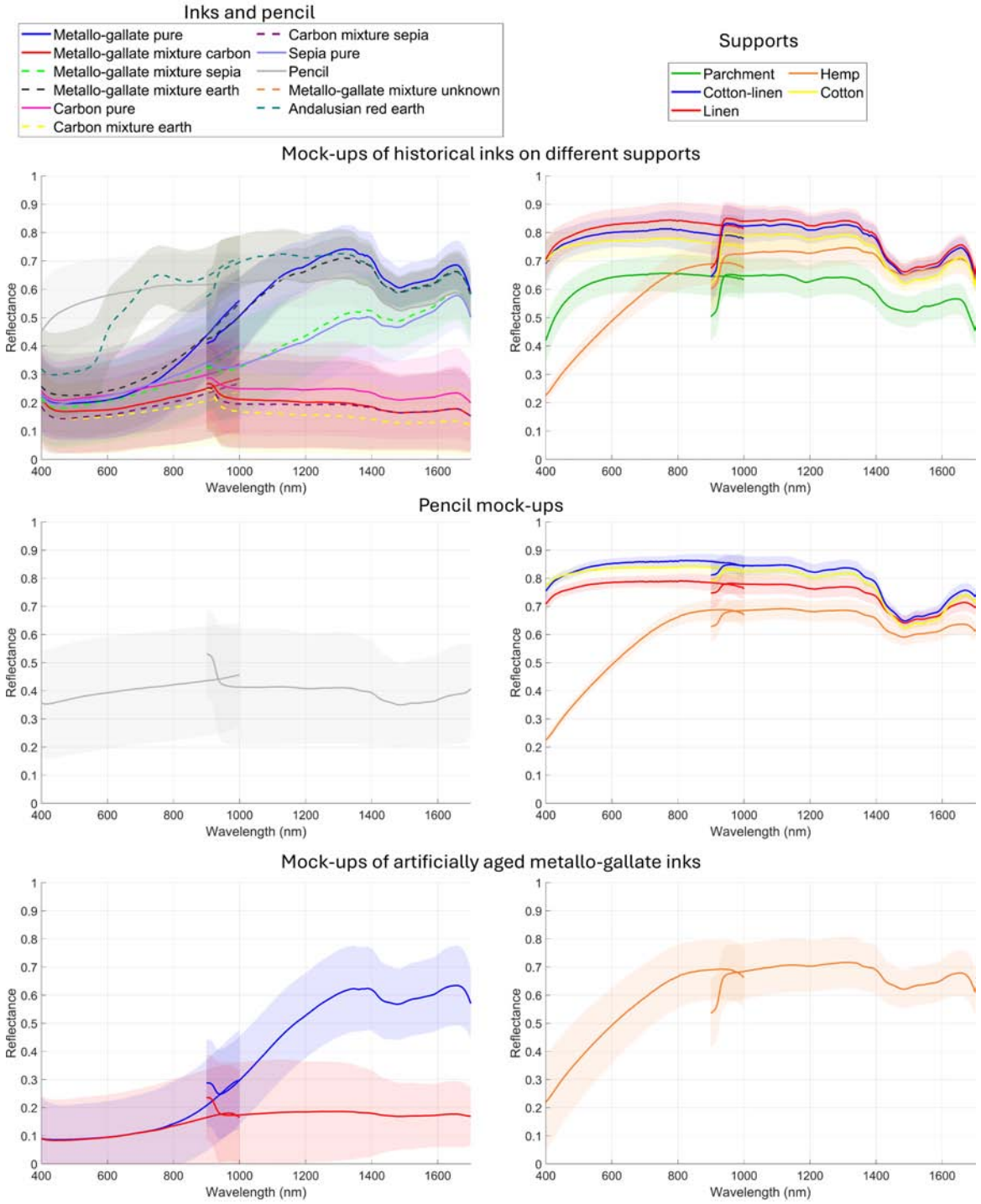

Figure 3.7: Average spectra and standard deviation per class (each line in the plot) and subset (each row) in the mock-up set. Inks and pencil are represented in the left graphs, while supports are shown in the right graphs.

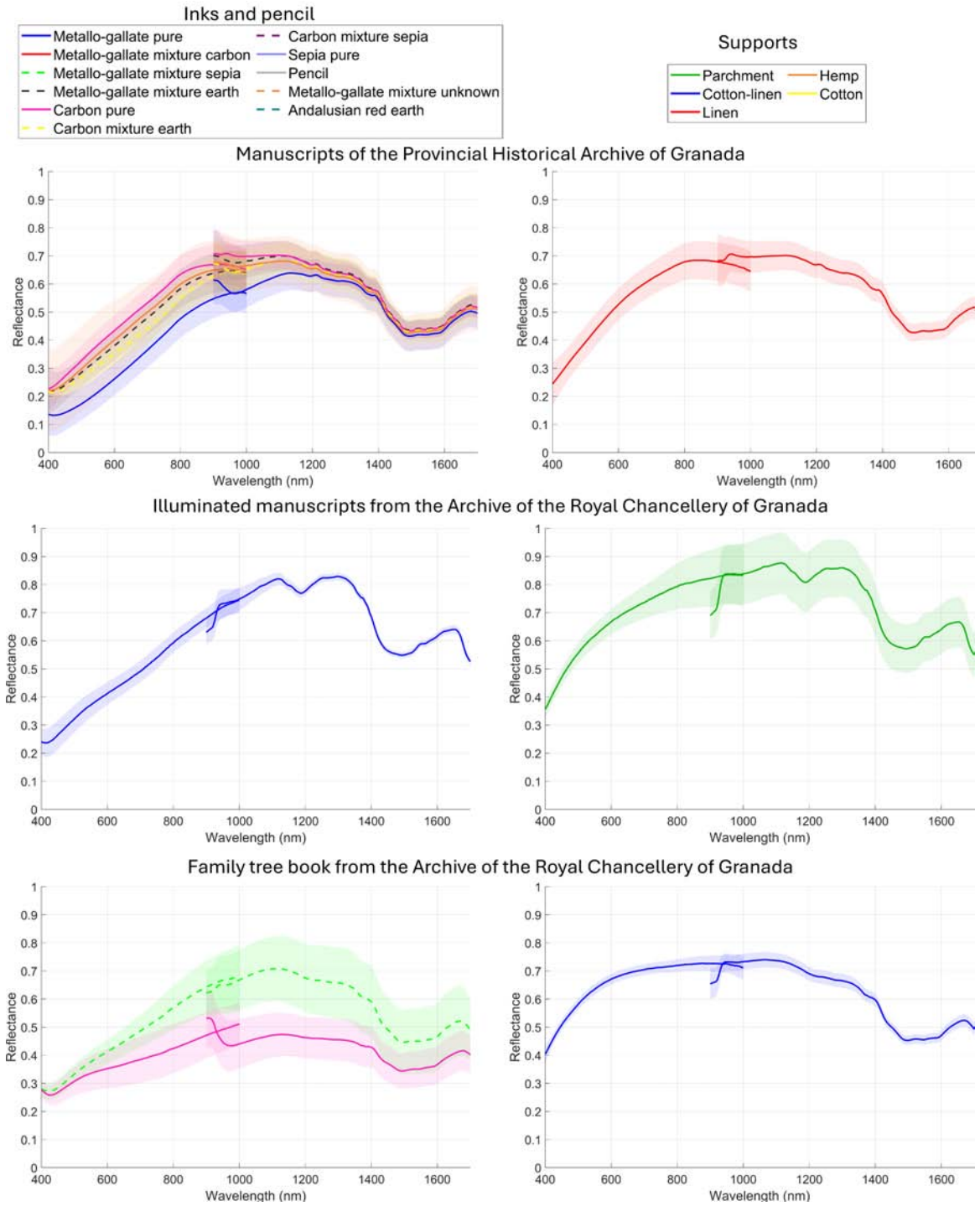

Figure 3.8: Average spectra and standard deviation per class (each line in the plot) and subset (each row) in the historical documents set. Inks are represented in the left graphs, while supports are shown in the right graphs.

duce reflectance, resulting in spectra resembling those of pure carbon-based inks. Pencil spectra generally exhibit flat shapes with lower reflectance values in the pencil mock-ups subset due to the inclusion of pencils with varying hardness, including very dark grades. In the subset of mock-ups with historical inks, only HB pencils were used. Andalusian red earth displays a spectrum characteristic of red hues, while it becomes transparent in the infrared.

Support materials such as cotton-linen, linen, and cotton exhibit similar reflectance spectra. Parchment shares a similar shape but demonstrates lower reflectance values, indicative of its darker tone. Additionally, parchment tends to show greater heterogeneity, particularly in historical documents, as modern parchment is generally more uniform in composition and appearance. Hemp has a different shape from the others, making it easily distinguishable from other supports.

For the historical documents, notable differences emerge when comparing spectra to those of mock-ups. For instance, carbon-based inks in historical manuscripts, particularly those from the Provincial Historical Archive of Granada (first row, Figure 3.8), do not exhibit the nearly complete absorption of infrared radiation seen in mock-ups. Instead, these inks also become partially transparent in the infrared range. This discrepancy could be attributed to aging processes, such as surface wear due to rubbing, which may reduce the ink layer thickness, causing the spectrum to resemble that of the underlying support. Similarly, minor variations in support spectra compared to mock-ups are likely due to aging effects.

Diversity in standard deviation is observed across subsets, derived from factors such as the use of different supports or variability within ink classes. For example, the pure metallo-gallate ink class includes inks with additives like pomegranate juice, myrtle infusion, or varying amounts of Cu, Zn, and Fe sulfates. Likewise, the pencil subset includes a range of hardness grades. This variability is intentional,

as it enhances the robustness of classification models by capturing a wide range of potential conditions, preparing them for real-world applications rather than highly controlled scenarios.

Comparing spectral libraries of black pigments or inks is challenging, as they consist of different samples prepared using diffrent techniques, binders, supports, and different data acquisition procedures or instruments. While it can be found only one HSI pigment database published in the range of 400 to 1000 nm [124], in general, the already existing libraries contain a single homogeneized measurement per sample made by using spectrorradiometers or FORS [123, 256]. As the database presented in this paper contains hundreds or thousands of datapoints, the average spectra of mock-up inks were used for comparison with existing ink spectral libraries.

The hyperspectral pigment database [124] includes seven pigments or inks that are also present in our dataset: grape black, ivory black, cherry black, bistre, atramentum, Andalusian red earth (or red ochre), and sepia. In our classification, the first four inks are grouped together as carbon-based inks due to the same origin and similar spectral characteristics. This grouping was validated by comparing the mean spectra provided by the authors of the database [257] with our data, confirming that carbon-based inks exhibit a consistently flat, low reflectance in the VNIR range. Similarly, the spectra for Andalusian red ochre, atramentum, and sepia show comparable shapes across both datasets.

Another publicly available database includes reflectance spectra obtained with spectrometers such as the GorgiasUV (200–1000 nm) and InGaAs (900–1700 nm) spectrometers [256]. This database contains seven black pigments or inks also present in our dataset: ivory black, vine black, bone black, lamp black, iron gall ink, Andalusian red ochre, and sepia. Again, the first four were grouped as carbon-based inks, exhibiting flat spectra between 400 and 1700 nm. The iron gall ink be-

comes transparent in the infrared, and it is interesting to see how the shape of the spectra in the 900-1700 nm range is completely influenced by the support, since in this case it was deposited on cardboard. Sepia ink becomes transparent beyond 1500 nm, displaying a common feature with the sepia spectra in our database, as shown in Figure 3.7 upper left plot.

Overall, our spectral data aligns well with existing databases of inks and black pigments, supporting the reliability of this dataset for further analysis and applications.

3.4.3 Principal Component Analysis (PCA)

PCA is a widely used dimensionality reduction technique in spectral data analysis, often employed for assessing the separability of datasets. For each minicube, the average spectrum was calculated for each class, and PCA was performed using these averaged spectra. Two principal components (PCs) were selected based on the Variance Accounted For (VAF) metric, identifying the inflection point where the VAF versus PCs curve flattens. PC1 explains 85.0% of the total variance, and PC2 explains 11.0%, resulting in a combined VAF of 96% with just two components. Figure 3.9 presents the score plots for PCs 1 and 2, showing inks and pencil in the left plot and supports in the right plot. Each class is represented by a unique color, and subsets are distinguished by different symbols.

In the left plot, the point clouds for pure metallo-gallate inks (dark blue) and their mixtures with earth (black) overlap, indicating that their spectra are highly similar. Similarly, pure carbon-based inks (pink) cluster closely with their mixtures with metallo-gallate ink (red), sepia (purple), or earth (yellow). These two groups, metallo-gallate-based and carbon-based inks, form distinct, separable clusters. However, the clouds for pure sepia (teal) and its mixture with metallo-gallate ink (light

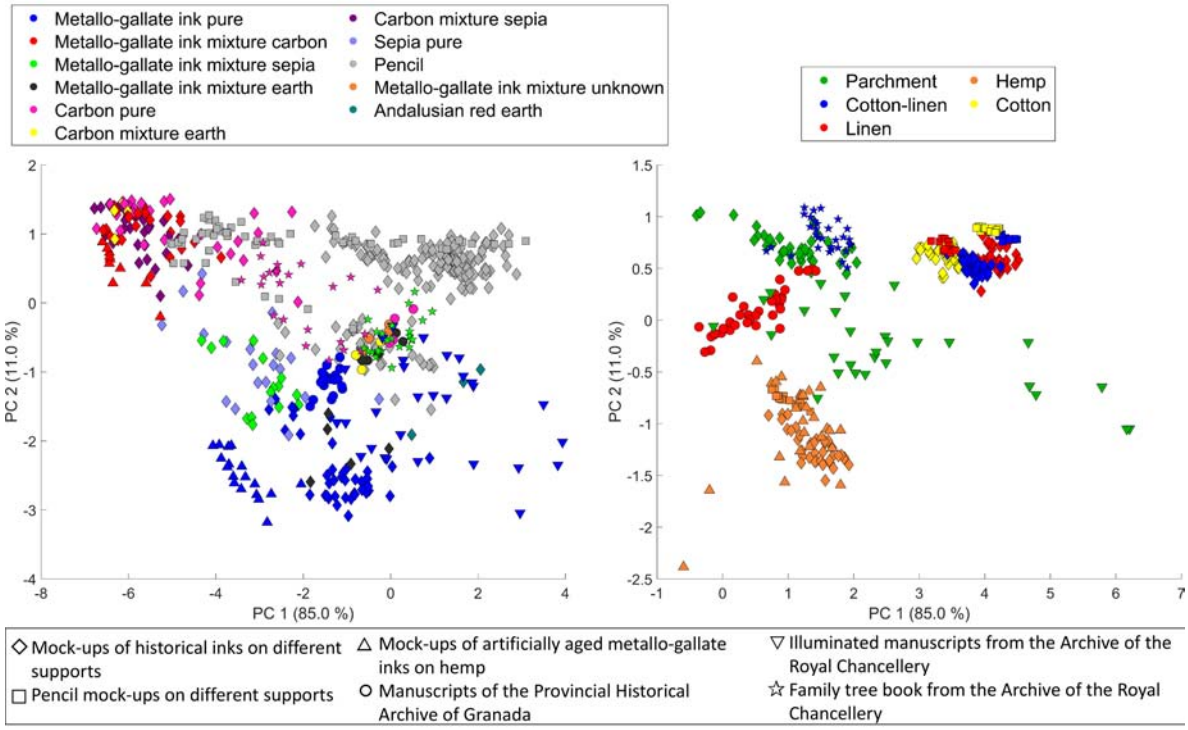

Figure 3.9: Score plots of PCA for inks and pencil (left) and supports (right). Colors indicate different classes, and symbols denote different subsets.

green) lie between these two clusters, reflecting intermediate spectral characteristics. Pencil samples, in contrast, form a more or less distinct cluster in a separate region. As for the standard deviation in the mean spectra plot, a high heterogeneity is found in the PCA due to different factors, including differences in spectra between mock-ups and historical documents, the application of the same inks on different supports, variations in ink recipes, and the grouping diverse inks within the same class. For example, the minicubes from the Manuscripts of the Provincial Historical Archive of Granada (depicted as circles) exhibit minimal variability, forming a concentrated cluster. A similar pattern is observed for the artificially aged mock-ups: they form two distinct clusters based on the type of ink and are separated from the non-aged mock-ups. In this case, the support does not influence these clusters, as all inks were applied on hemp paper. However, samples from other subsets and classes are more

dispersed, often overlapping in the PCA space.

In the right plot, hemp samples are clearly separated from other supports, forming a distinct cluster. However, notable differences can be observed between mock-ups and historical document samples of the same support types. For instance, cotton-linen samples (blue clouds) and linen samples (red clouds) from mock-ups and historical documents form separate clusters despite belonging to the same class. Parchment samples (green clouds) are distributed across a wider area, also showing a clear distinction between mock-ups and historical documents. In contrast, cotton samples are tightly grouped, reflecting the limited variability within this class, as there were no samples with pure cotton support in the historical documents. Overall, the results of the PCA analysis show that separability among classes is not enough for tackling material identification using PCA components as input. They also suggest that the database is wide enough to cover for a fair amount of the variability found in both mock-ups and historical document samples.

3.5 Usage Notes

Example code, available on GitHub [244], is provided in both MATLAB and Python to perform the following tasks:

- Exploration of the general content of the minicube stored as an HDF5 file, including access to metadata stored as attributes.
- Extraction of the HSI data from the dataset named `DataCube`, along with relevant attributes as variables in the workspace or environment.
- Visualization of false-color RGB images derived from selected spectral bands.
- Extraction of the GT data.

Figure 3.10: False-color RGB images (left) created using bands [645, 565, 440] nm (top) and [1700, 1300, 1100] nm (bottom), GT image (center), and the mean spectral reflectances with standard deviation for the labels present in the GT, extracted from the minicubes '00007-VNIR-mock-up.h5' and '00007-SWIR-mock-up.h5' (right).

- Calculation of the mean and standard deviation for each class in the GT using the hyperspectral data from the `DataCube`.
- Plotting of the mean reflectances and their standard deviations for both the VNIR and SWIR minicubes.

To execute the MATLAB code, a MATLAB version R2011a or later is required, along with the Image Processing Toolbox. For Python, the following packages are needed: `h5py`, `numpy`, `pillow`, and `matplotlib.pyplot`.

The resulting images based on the provided code are shown in Figure 3.10, using the minicubes '00007-VNIR-mock-up.h5' and '00007-SWIR-mock-up.h5'.

For the complete HYPERDOC database of 1681 samples, which is available online on the Color Imaging Lab group's website [225], a user-friendly visualization tool has been developed in Python. This tool is currently under development for the inclusion of more options related to material identification and will be released in its final form in the near future.

3.6 Conclusions, limitations, and future perspectives

One limitation of the HYPERDOC database is the imbalance between the mock-up samples and those derived from historical documents. Mock-ups offer a controlled and reproducible framework ideal for training robust algorithms and investigating specific interactions between inks and supports. Simultaneously, the inclusion of historical samples provides valuable insights into real-world conditions of aging and degradation, broadening the applicability of the database to practical scenarios. Efforts have been made to include a broader range of historical materials; however, access to such documents remains a challenge due to their fragile nature and restricted availability. Moreover, the aging processes of inks and supports, along with their interactions and the uncertainty about the recipes used to prepare inks and supports in historical documents introduce further variability in the spectral properties of these materials, which are challenging to model. These processes may result in spectral features that diverge from those of mock-ups, potentially reducing the database's generalizability to other historical samples of different periods and origins. The complete HYPERDOC database [225] includes additional historical samples to address this limitation. However, a drawback is that GT data and/or material information are not available for all the samples.

Additionally, identifying the materials present in historical samples often requires complementary analytical techniques. Without these methods, it is difficult to precisely determine the composition of certain inks, pigments, or supports, which can affect the accuracy of the GT annotations.

The GT annotations in this database were generated using a semi-automatic approach, including manual corrections, which ensures a reasonable degree of accuracy. However, human involvement may still introduce occasional inconsistencies,

particularly when defining boundaries between inks and supports. Interactions between inks and supports, such as penetration or blending at the interface, create transitional zones where both materials are mixed, complicating precise annotation. To address this challenge, algorithms could be employed to erode or expand the annotated ink regions in the GT.

Future efforts should focus on addressing these limitations by expanding the database to include more historical samples with greater material variability and developing more robust and automatic methods for GT annotation that minimize human error, although this matter is still unsolved in the literature so far.

Despite its limitations, the HYPERDOC database is a unique and comprehensive resource that integrates mock-ups and historical documents, supporting diverse applications as previously demonstrated. It bridges the gap between the scientific and conservation communities, promoting the adoption of advanced techniques such as HSI, which remain relatively novel in the field. By fostering interdisciplinary collaboration and enabling the development of innovative methodologies, HYPERDOC contributes to the analysis of historical documents, ensuring its ongoing relevance and advancing the safeguarding of our CH.

3.7 Code Availability

Hyperspectral data capture and reflectance correction were performed using Spectronon Pro 3.5.5. MATLAB code for converting from BIL to MAT format, from MAT to HDF5, and for extracting GT data, along with additional MATLAB and Python code for the visualization and analysis of minicubes, is available on GitHub: [244].

CHAPTER 4

Selection of optimal spectral metrics for classification of inks in historical documents using HSI data

In this chapter, three methods for selecting optimal spectral metrics for classifying historical black inks using HSI data are presented. It is the result of specific objective O3 of this PhD thesis.

This study has been published as:

A. B. López-Baldomero*, M. Martínez-Domingo, E. M. Valero, R. Fernández-Gualda, A. López-Montes, R. Blanc-García, and T. Espejo, "Selection of optimal spectral metrics for classification of inks in historical documents using hyperspectral imaging data," in *Optics for Arts, Architecture, and Archaeology (O3A) IX*, vol. 12620, pp. 99–111, SPIE, 2023.

4.1 Introduction

As stated in chapter 1, material characterization provides valuable information, and specifically ink analysis serves as an essential tool for codicologists and historians, who explore both the content and material composition of manuscripts to gather this information [31]. In general, the study of inks provides significant insights into the sociocultural and technological changes in historical document production [30].

Previous studies have evaluated different spectral metrics, especially in the context of pigment identification in artworks (as detailed in subsubsection 2.5.1.1). Some research has focused on identifying modern ballpoint pen inks [129,223,258], particularly in the context of forensic analysis [221,222,259]. However, prior to our study, the use of HSI for identifying inks in ancient manuscripts remained unexplored, including the application and evaluation of spectral similarity metrics for this purpose.

Three methods were used for selecting optimal spectral metrics for the purpose

of classifying inks of historical interest using HSI in the VNIR and SWIR spectral ranges. The metrics tested were $RMSE$ [260], SAM [261], SID [262], $SIDSAM$ [263], Normalized Spectral Similarity Score ($NS3$) [264], and $JMSAM$ [265]. Initially, a performance comparison of these spectral metrics was conducted using mock-up samples to determine the most suitable ones. The best three metrics were used to classify historical samples of ancient manuscripts.

This work represents the first attempt to apply spectral imaging for the automatic classification of inks in historical documents, serving as an initial approach to the problem. It adopts a simplified perspective, using spectral metrics to conduct a preliminary study of the problem's complexity.

4.2 Methodology

4.2.1 Hyperspectral image capture

Hyperspectral images of mock-ups and historical samples were acquired in the VNIR and SWIR ranges following the capture process and post-processing described in subsection 3.2.2 and subsubsection 3.2.3.1.

4.2.2 Mock-up samples and reference libraries

Eight mock-up samples of written words were made using six pure inks plus two mixtures, applied on Canson® watercolor paper. These samples are part of the complete HYPERDOC database but are not included in chapter 3 as they were not spatially registered. The pure inks used in the study were Iron Gall Ink (IGINK), Vine Black (BVINE), Ivory Black (BIVOR), Bone Black (BBONE), Lamp Black (BLAMP), and Sepia (SEPIA), while the mixtures used were Lamp black with Iron Gall Ink (IG-

ILA), and Sepia with Iron Gall Ink (IGISE). From now on, the acronyms in brackets will be used for brevity. All the inks used in this study were self-made according to ancient recipes [236, 237].

Two spectral reflectance libraries were obtained, one for each spectral range, including these eight inks. The libraries were generated by averaging the spectral reflectances of pixels within the central zone of the strokes, extracted using the *imerode* function implemented in MATLAB®. This approach aimed to minimize the influence of the support in the spectra.

The values of the spectral metrics in the mock-up samples were then studied to determine the classification confidence threshold for each metric. This was achieved by comparing pixel-wise the entire strokes with the reference libraries, and using bar graphs to determine the similarity metric threshold value for which a reliable classification could be achieved. In addition, the score maps, i.e. the spectral metrics value spatial distributions obtained comparing each pixel of the image with the reference library, were calculated. In total, eight score maps were obtained for each metric studied, corresponding to each cube of the mock-up samples.

4.2.3 Spectral similarity metrics

Spectral similarity metrics quantify the degree of spectral similarity between two spectra. In HSI, these metrics play a crucial role in calculating the similarity between each pixel in an image and a reference spectrum associated with a particular material. In the context of CH, the selection of spectral metrics is significant as they need to be sensitive to variations in the magnitude of spectra, especially when dealing with mixed or aged pigments [126]. Six similarity metrics have been studied, which are described below. It is important to note that lower values of these metrics indicate a higher degree of spectral similarity between the reference and test spectra.

Additionally, all metrics are calculated pixelwise, comparing each pixel's spectrum in the hyperspectral image, $t(x, y, \lambda)$, to the reference spectra in the library, $r(\lambda)$.

4.2.3.1 RMSE

The Root Mean Square Error ($RMSE$) is based on a Euclidean measure of the error between the reference, r_i, and the test spectrum, t_i (see Equation 4.2.1) [260].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - r_i)^2} \tag{4.2.1}$$

Where t_i represents the value of the test spectrum at the ith spectral band, and r_i corresponds to the value of the reference spectrum at the ith spectral band. N denotes the number of spectral bands.

4.2.3.2 SAM

The Spectral Angle Mapper (SAM) has been widely used in the CH domain for pigment and ink classification [129,130,266], due to its fastness and ease of use. This metric measures the angle between two spectra, treated as N-dimensional vectors in hyperspace, being N the number of spectral bands [261] (see Equation 4.2.2). This angle does not change with the modulus of the vectors, thus being insensitive to magnitude shifts in the spectrum.

$$\alpha = cos^{-1}\left(\frac{\sum_{i=1}^{N} t_i r_i}{\sqrt{\sum_{i=1}^{N} t_i^2}\sqrt{\sum_{i=1}^{N} r_i^2}}\right) \tag{4.2.2}$$

In Equation 4.2.2, r_i is the reference spectrum at the ith spectral band, t_i is the test spectrum at the ith spectral band, N is the number of bands, and α is the spectral angle in radians.

4.2.3.3 SID

Spectral Information Divergence (SID) measures the similarity between two spectra based on the probabilistic discrepancy between them [262]. The probability distribution of both spectra studied is expressed as:

$$p_i = \frac{t_i}{\sum_{i=1}^{N} t_i} \tag{4.2.3}$$

$$q_i = \frac{r_i}{\sum_{i=1}^{N} r_i} \tag{4.2.4}$$

Being r_i the reference spectrum at the ith spectral band, t_i the test spectrum at the ith spectral band, and N the number of bands. SID is then calculated as:

$$SID = \sum_{i=1}^{N} p_i log\Big(\frac{p_i}{q_i}\Big) + \sum_{i=1}^{N} q_i log\Big(\frac{q_i}{p_i}\Big) \tag{4.2.5}$$

4.2.3.4 SIDSAM

This metric merges as a combination of SID and SAM. There are two methods to obtain this metric (Equation 4.2.6 and Equation 4.2.7):

$$SIDSAM = SID \cdot tan(SAM) \tag{4.2.6}$$

$$SIDSAM = SID \cdot sin(SAM) \tag{4.2.7}$$

$SIDSAM$ calculates the perpendicular distance between the test and reference vectors or spectra [263]. It improves the spectral discriminability by making dissimilar spectra more distinctive and similar spectra more comparable. It has been used

to assess spectral characteristics of dye- and pigment-based ink prints [267], and for modern ink classification [129].

4.2.3.5 NS3

The Normalized Spectral Similarity Score ($NS3$) [264] combines both spectral angle and amplitude information, enhancing the effectiveness and accuracy of searching for matches within the reference library. This method is based on the Euclidean and SAM distances between two spectra (Equation 4.2.9). First, the Euclidean distance is computed (Equation 4.2.8).

$$A_{Euclidean} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - r_i)^2} \tag{4.2.8}$$

Being r_i the reference spectrum at the ith spectral band, t_i the test spectrum at the ith spectral band, and N the number of bands. The spectral angle (SAM) in radians α is computed as in Equation 4.2.2, and $NS3$ is calculated as:

$$NS3 = \sqrt{A^2_{Euclidean} + (1 - cos\alpha)^2} = \sqrt{A^2_{Euclidean} + cGFC^2} \tag{4.2.9}$$

Being $cGFC$ the complement of the Goodness-of-Fit Coefficient, a spectral metric that will be further used in chapter 6 and chapter 7.

4.2.3.6 JMSAM

The Jeffries-Matusita (JM) [265] method is a pairwise distance primarily employed in binary classification tasks. Several extensions of JM have been proposed for its use in multiclass classification [268]. The most common approach involves computing the average JM distance for all pairs of classes [269]. This metric has been applied to analyze dye- and pigment-based prints [267], and for modern ink

classification [129]. In this metric, the JM and SAM algorithms are orthogonally projected by using either a sine (Equation 4.2.10) or a tangent (Equation 4.2.11) function.

$$JMSAM = JMD \cdot sin(SAM) \tag{4.2.10}$$

$$JMSAM = JMD \cdot tan(SAM) \tag{4.2.11}$$

Being JMD the Jeffries-Matusita distance, which can be calculated as:

$$JMD = 2(1 - e^{-B}) \tag{4.2.12}$$

$$B = \frac{1}{8}(\mu_t - \mu_r)^T \left[\frac{\sigma_t + \sigma_r}{2}\right]^{-1} (\mu_t - \mu_r) + \frac{1}{2} ln \left[\frac{\left|\frac{\sigma_t + \sigma_r}{2}\right|}{\sqrt{|\sigma_t|\,|\sigma_r|}}\right] \tag{4.2.13}$$

μ_t and μ_r are the mean of the test and reference spectra, σ_t and σ_r are the covariance of the test and reference spectra.

4.2.4 Selection of optimal metrics for classification

Three methods are tested to determine the optimal spectral similarity metrics.

4.2.4.1 Method 1: Study of the confidence thresholds for each metric

After obtaining the spectral metric values by comparing the spectral reference libraries with the mock-up samples on a pixelwise basis, confidence thresholds were established for each metric. These thresholds were determined by computing the average metric values when comparing the samples with the inks present on them across all mock-up samples, taking the mean across all studied samples. A metric

value below the confidence threshold indicates a reliable classification. Based on this threshold, the percentage of pixels falling below these values was calculated. The optimal metric was selected as the one yielding the highest percentage of pixels below the confidence threshold.

4.2.4.2 Method 2: Evaluation of the classification

From the spectral metric values obtained comparing our spectral libraries with the mock-up samples, the classification maps were generated by labeling each pixel in the images with the material from the spectral library that gives the lowest metric value (maximum similarity criteria). This was done for all eight mock-up samples, considering the six spectral metrics and the two ranges studied. Then, the classification performance was evaluated using the accuracy and F1-score, calculated as detailed in subsubsection 2.5.1.2. The mean and standard deviation of both evaluation metrics were obtained for all the samples studied and each of the six spectral metrics. From these results, the spectral metrics providing the highest values of these evaluation metrics were selected.

4.2.4.3 Method 3: Estimation of the Probability of Spectral Discrimination (PSD) and the Power of Spectral Discrimination (PWSD)

In a third approach, the Probability of Spectral Discrimination (PSD) and the Power of Spectral Discrimination ($PWSD$) were calculated to determine the optimal spectral metric [270, 271].

The PSD calculates the probability for all spectra, $\{s_j\}_{j=1}^{J}$, in a reference library to classify a test spectrum, t, as:

$$PSD = \frac{m(t, s_i)}{\sum_{j=1}^{J} m(t, s_j)} \quad \text{for } i = 1, \ldots, J \tag{4.2.14}$$

Where $\sum_{j=1}^{J} m(t, s_j)$ is the normalization constant and $m(,)$ is any of the previously defined spectral similarity metrics. The PSD was calculated comparing all the spectra in the reference library. The higher the PSD, the easier to discriminate one spectrum from another. This allows the comparison of spectral metrics in terms of their ability discriminating the spectra in the libraries.

The $PWSD$ evaluates the effectiveness of spectral similarity metrics. It is formulated based on the capacity to differentiate one pixel from another compared to a reference pixel (spectra), denoted as r. The $PWSD$ quantifies the effectiveness of distinguishing between two spectra, represented by s_i and s_j, in relation to the reference spectra r. The spectral discrimination power is defined as follows:

$$PWSD = max \left\{ \frac{m(s_i, r)}{m(s_j, r)}, \frac{m(s_j, r)}{m(s_i, r)} \right\} \tag{4.2.15}$$

$PWSD$ equals 1 if $s_i = s_j$ and, for certain spectral metrics, when both spectra are equidistant from r. For two distinct spectra at different distances from r, $PWSD$ cannot be less than 1, as one spectrum will always be closer to r than the other, ensuring that one of the compared ratios exceeds 1 (the maximum), while the other remains below. Consequently, values closer to 1 indicate a lower discriminative ability of the spectral similarity metric for distinguishing r based on the two spectra. By analyzing $PWSD$, the degree of spectral differentiation can be quantified, helping to identify potential classification ambiguities.

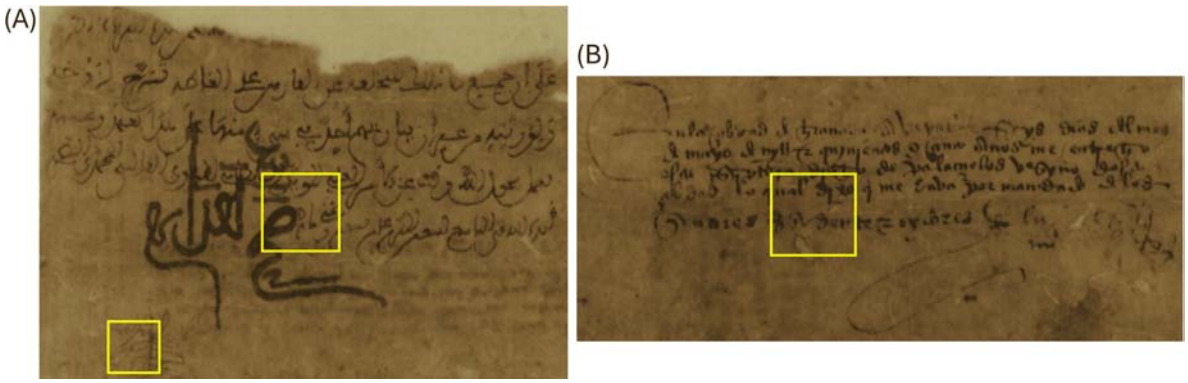

Figure 4.1: False RGB images (645, 565, 440 nm) from the VNIR capture of the manuscripts, highlighting minicube extraction areas: (A) deed of sale for a shop, (B) restitution of properties.

4.2.5 Historical samples

Six minicubes (three in the VNIR range and three in the SWIR) extracted from two 15th-century manuscripts were used to classify historical inks by searching through spectral libraries with optimal spectral metrics. These minicubes are part of the complete HYPERDOC database but were not included in chapter 3 as they were not spatially registered. The reason is that at the time this research began (see Figure 1.3 for reference), the database was still in its early stages.

The manuscripts, known as *Nasrid ʻAlāmas*, originate from the Archive of the Royal Chancellery of Granada and date to 1461 and 1485. The *ʻalāma* refers to a short phrase written in official documents of medieval and early modern Arabic chancelleries, that served both as validation and as a symbolic representation of the sultan issuing the document.

The first manuscript analyzed is a deed of sale for a shop in the Granada souk (Figure 4.1 (A)), while the second is a restitution of properties to the heirs of a deceased nobleman (Figure 4.1 (B)). The inks used in these documents consist of a mixture of iron gall ink and an unidentified organic component, applied on parch-

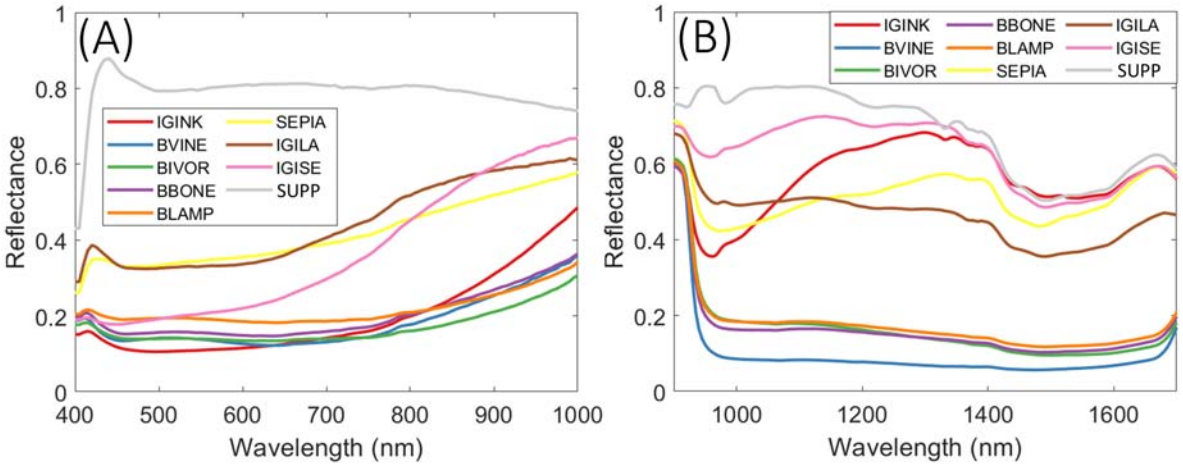

Figure 4.2: Spectral reflectances from the reference libraries and the support in the (A) VNIR range, and (B) SWIR range.

ment and Arabic paper [22]. The support material was identified using optical microscopy and high-resolution Variable Pressure Scanning Electron Microscopy (VPSEM), while VPSEM was also employed to analyze the inks. However, as this study is still ongoing, the specific organic component in the mixed ink has not yet been determined.

4.3 Results

In Figure 4.2, the spectral reflectances of the reference libraries and the support are shown for both VNIR and SWIR ranges. It can be seen that spectral reflectance of carbon-based inks (BVINE, BIVOR, BBONE and BLAMP) have a similar shape in both ranges. Iron gall ink, sepia and mixtures are clearly separated from carbon-based inks in the SWIR range, which a priori could be an advantage for the classification.

The results for the three proposed methods for selecting optimal spectral metrics are presented in the following subsections.

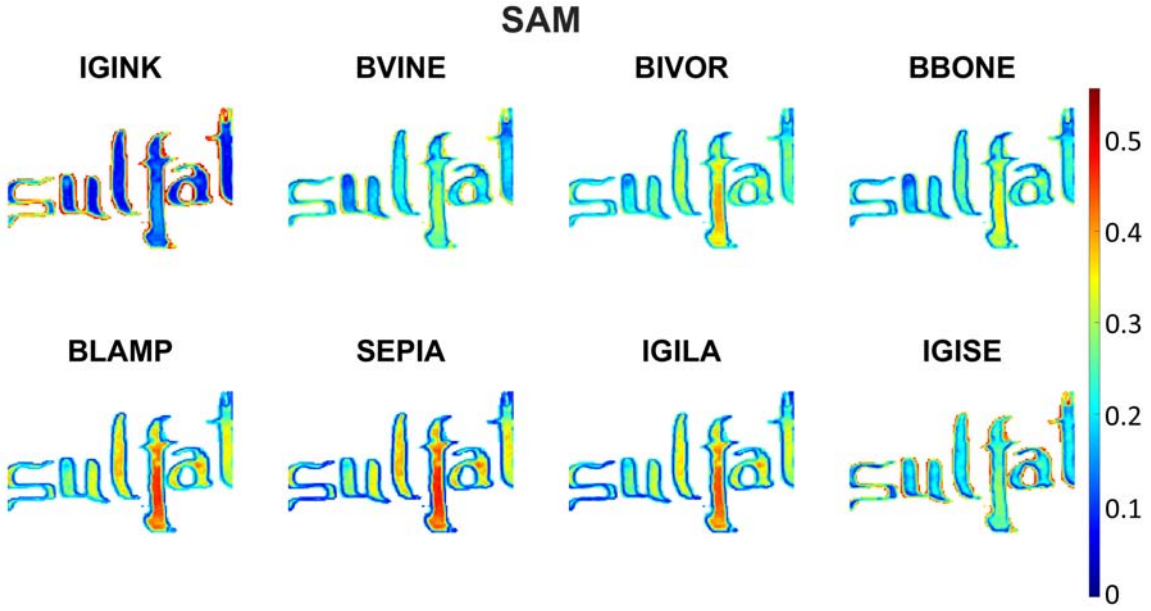

Figure 4.3: Score maps obtained with the SAM metric for the iron gall ink sample compared with the eight inks of the library in the VNIR range.

4.3.1 Method 1: Study of the confidence thresholds for each metric

As an instance, the score maps obtained using the SAM metric for the iron gall ink in the VNIR range are shown in Figure 4.3. The lowest value of SAM is achieved when comparing the cube with the IGINK of the reference library, which is the only ink present in the sample. In addition, it is interesting to highlight the effect of the stroke edges, which tend to have higher metric values. This could be due to the interaction of the ink and the support.

In order to study the score maps for each metric and spectral range, we calculated the average and standard deviation when comparing the ink present in the cube with the correct ink of the reference library (i_p), as well as with the other inks not present in the cube (i_{np}). This analysis was performed for all eight mock-up samples and the results are represented in bar graphs. The results for $JMSAM$ and $SIDSAM$ metrics, which provided the highest and lowest variances, are shown in Figure 4.4.

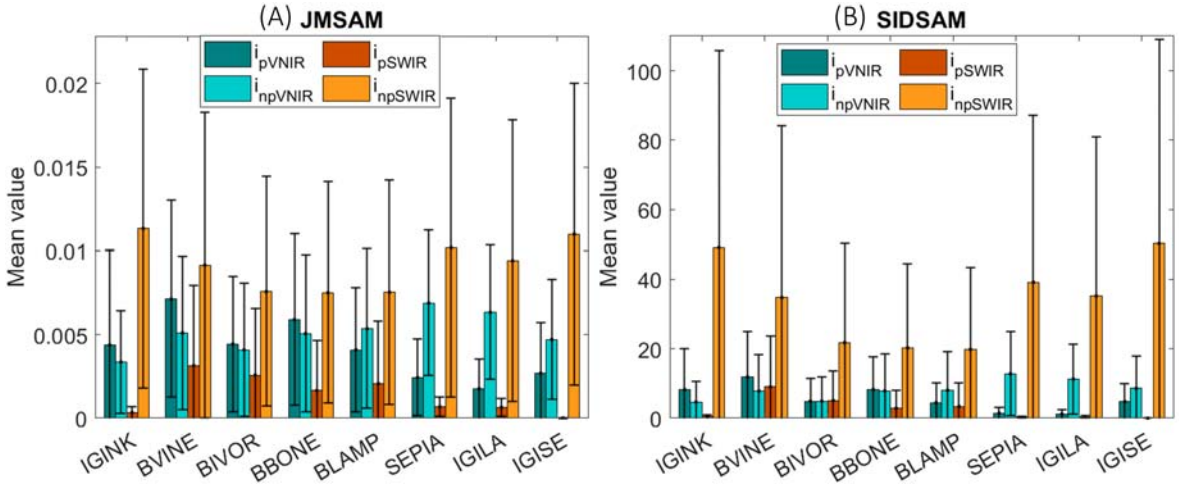

Figure 4.4: Mean (bars) and standard deviation (lines) of metrics values in the VNIR and SWIR ranges for (A) $JMSAM$, and (B) $SIDSAM$, comparing the ink present in the cube with the correct ink in the reference library (i_p) and the incorrect inks (i_{np}).

By examining the bar graphs, we can compare the behavior of different inks in both spectral ranges. In the SWIR range, the IGINK and the mixture with sepia (IGISE) exhibit the greatest dissimilarity when compared to the other inks. This indicates that they would be relatively easier to distinguish from the rest. Conversely, BIVOR would present challenges in differentiation from the other inks. In the VNIR range, distinguishing BIVOR and BBONE from the rest of inks would be particularly challenging, whereas the mixture IGILA would be more readily recognizable. It is interesting to note that in the VNIR range, for some inks like BVINE and IGINK, the metrics values obtained when comparing the cube pixels with the ink that is present are higher than when comparing them with the inks that are not present. This suggests that differentiating these inks from the others in the library can be quite challenging.

From these results, the average of the metrics values when comparing with the inks present in the cubes was calculated for all the mock-up samples (mean of the dark green and dark orange bars in Figure 4.4) and established as the confidence

Table 4.1: Confidence thresholds for the six metrics in both spectral ranges.

	RMSE	**SAM**	**SID**	**SIDSAM**	**NS3**	**JMSAM**
VNIR	0.1806	0.1993	17.5593	5.6627	0.1831	0.0041
SWIR	0.0869	0.1500	7.1928	2.6943	0.0911	0.0014

Table 4.2: Percentage of pixels with metrics values lower than the confidence threshold. Mean for all the mock-up samples.

	RMSE	**SAM**	**SID**	**SIDSAM**	**NS3**	**JMSAM**
VNIR	78.93	83.61	96.13	**99.47**	79.00	90.12
SWIR	86.26	83.91	92.22	**98.99**	86.99	94.02

thresholds (shown in Table 4.1). A metric value under these confidence thresholds indicates a reliable classification. As shown in Figure 4.4, these thresholds would be effective for certain inks but may not be suitable for all of them.

For the mock-up samples, the percentage of pixels with metrics values lower than the confidence threshold indicated in Table 4.1 were calculated (Table 4.2). A higher percentage indicates better performance.

From Table 4.2, we can see that $SIDSAM$ outperformed the other metrics in both spectral ranges. $RMSE$ and SAM exhibited the lowest percentage of pixels under the confidence thresholds for the VNIR and the SWIR range respectively, closely followed by NS3.

4.3.2 Method 2: Evaluation of the classification

From the score maps, the classification maps were obtained. In Figure 4.5 the classification maps for the iron gall ink in the SWIR range are shown for the six spectral metrics studied. Once again, the effect of the edge of the strokes can be observed in this Figure. For all the metrics studied the central area of the stroke is correctly classified as IGINK, while the edges are classified as a mixture with sepia (IGISE). This phenomenon is observed consistently across the other mock-up samples as well.

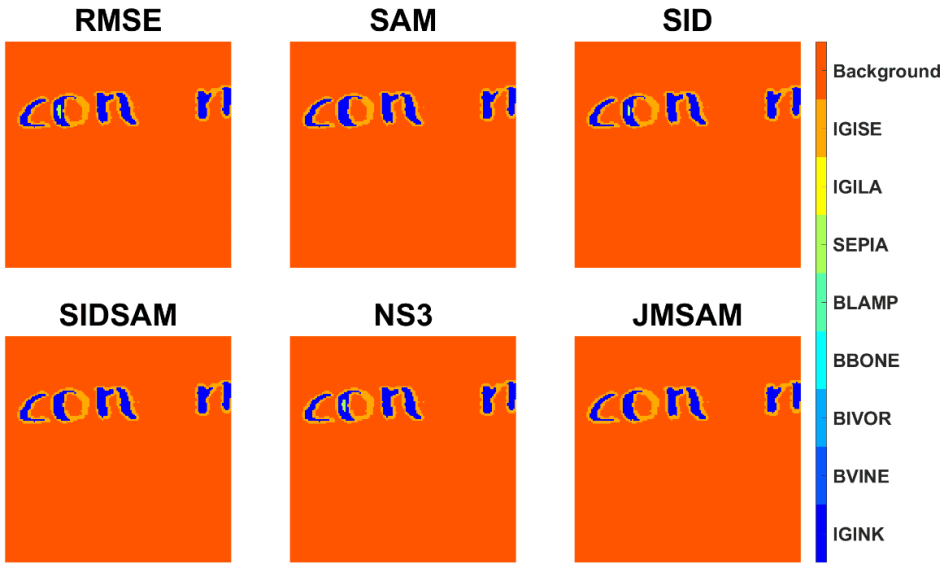

Figure 4.5: Classification maps for the iron gall ink sample in the SWIR range.

Table 4.3: Mean (standard deviation) of accuracy and F1-score for all the inks studied and the six spectral metrics.

		RMSE	SAM	SID	SIDSAM	NS3	JMSAM
VNIR	Accuracy	0.4191 (0.2650)	**0.4442 (0.2520)**	0.4147 (0.2627)	0.4282 (0.2587)	0.4235 (0.2650)	0.3903 (0.2359)
	F1-score	0.5550 (0.2653)	**0.5814 (0.2625)**	0.5497 (0.2601)	0.5652 (0.2480)	0.5610 (0.2644)	0.5225 (0.2467)
SWIR	Accuracy	0.2871 (0.2606)	**0.3210 (0.2553)**	0.2857 (0.2626)	0.3110 (0.2567)	0.2928 (0.2562)	0.3040 (0.2680)
	F1-score	0.3991 (0.2271)	**0.4427 (0.2242)**	0.3987 (0.2301)	0.4323 (0.2251)	0.4064 (0.2208)	0.4276 (0.2354)

The mean and standard deviation of accuracy and F1-score (computed as described in subsubsection 2.5.1.2) for the classification of all mock-up samples using both libraries are shown in Table 4.3 for each spectral metric. From Table 4.3, it can be observed that the metric that provided the highest accuracy and F1-score for both spectral ranges was SAM. This is consistent with the results found by Devassy *et al.* [129]. They conducted a comparative analysis of five similarity metrics to assess their ink classification capabilities: SAM, SCM, ED, SID, and BE, obtaining a

higher accuracy with the SAM metric. In addition, the VNIR range showed higher values of accuracy and F1-score compared to the SWIR range, which could be unexpected looking at the bar graphs of Figure 4.4. This difference in performance could be attributed to the fact that in the SWIR range, certain inks such as IGINK and SEPIA become transparent, while carbon-based ink do not. As a result, the spectral difference between these two groups is more pronounced. Nevertheless, when it comes to classification, the spectra of those inks which are transparent in SWIR are very similar because they are influenced by the support, making them more difficult to discriminate both from each other and from the support itself. This can be observed by comparing the spectrum of the support (SUPP) with those of the inks (IGINK, IGISE, IGILA, and SEPIA) in Figure 4.2, particularly in the 1400-1700 nm range. The same applies to carbon-based inks, which have similar spectra. In general, the classification results show relatively low accuracy, with performance around 50% or even lower across all metrics. Several factors may contribute to this. First, the borders of the strokes present a challenge, as variations in the reflectance spectra in these areas make classification more difficult compared to the central part of the strokes. Another factor is the simplicity of the method used, as we are relying on spectral similarity metrics. Additionally, differentiating between a large group of inks, making it a multiclass problem with a total of nine classes, is challenging, particularly when some inks are spectrally very similar, such as BVINE, BIVOR, BBONE, and BLAMP, all carbon-based inks, as shown in Figure 4.2. These preliminary results helped guide our approach to the future classification work, presented in chapter 5, where we grouped the different classes.

Table 4.4: Ratio of PSD values between the inks not present in the sample and the ink that is present.

	Ink	RMSE	SAM	SID	SIDSAM	NS3	JMSAM
VNIR	IGINK	**6.16**	3.12	4.47	1.92	**6.16**	3.73
	BVINE	5.21	**5.87**	3.97	3.77	5.21	4.09
	BIVOR	6.04	**9.58**	4.27	6.85	6.04	5.36
	BBONE	5.89	**7.39**	5.44	6.63	5.89	5.40
	BLAMP	7.10	13.38	8.10	**17.66**	7.10	9.55
	SEPIA	15.52	16.32	51.39	**128.46**	15.52	32.68
	IGILA	18.49	12.93	69.76	**142.44**	18.49	53.17
	IGISE	11.10	4.58	**20.68**	17.43	11.10	20.12
SWIR	IGINK	35.92	31.27	404.03	**3359.76**	35.92	635.55
	BVINE	17.35	5.74	25.71	32.61	17.35	**39.11**
	BIVOR	18.07	9.93	31.62	52.75	18.07	**52.97**
	BBONE	23.65	13.69	58.08	**138.61**	23.65	115.88
	BLAMP	4.09	4.90	4.22	**6.44**	4.09	4.20
	SEPIA	27.11	36.80	263.71	**2556.92**	27.11	226.05
	IGILA	20.46	45.95	143.29	**1639.72**	20.46	198.47
	IGISE	104.91	101.11	4015.00	**100515.16**	104.91	12077.51

4.3.3 Method 3: Estimation of the PSD and the PWSD

Table 4.4 presents the ratio between the PSD values of the inks not present in the sample and the PSD value of the ink that is actually present. A higher value in this division indicates a greater probability of spectral discrimination between the ink present in the sample and the other inks. The metric providing the highest value for each ink has been highlighted in bold.

The results shown in Table 4.4 indicate that depending on the ink being classified, some metrics will provide a higher probability of discrimination than others. In general, higher probabilities are obtained in the SWIR spectral range, particularly with the $SIDSAM$ metric which provided significantly higher values for inks such as SEPIA and the two mixtures. In this range, the $JMSAM$ and $SIDSAM$ metrics provided the best values. In the VNIR range, some carbon-based inks showed the

best results with SAM, while $SIDSAM$ provided the best values for other three inks. From the evaluation of the $PWSD$, the results showed that in the VNIR range the metric that provided the highest $PWSD$ was $JMSAM$, followed by $SIDSAM$ and SID. $RMSE$ and $NS3$ metrics performed the worst in terms of $PWSD$ in this range. In the SWIR range, the $SIDSAM$ metric outperformed all the other metrics, providing 5 times better results than the next best metric, which was $JMSAM$. Again, $RMSE$ and $NS3$ provided the worst results. The detailed results are not presented here due to the large volume of data generated, as each spectrum in the spectral library was compared with all other spectra in the library and with each of the studied minicubes, across every metric and both spectral ranges.

Summarizing the results of the PSD and $PWSD$ metric assessment methods, $SIDSAM$ and $JMSAM$ were selected as optimal metrics.

4.3.4 Optimal metrics applied to historical samples

The optimal spectral metrics, selected according to the three approaches discussed above, were applied to three fragments of historical documents. The selected metrics were $SIDSAM$ from method 1 and 3, SAM from method 2, and $JMSAM$ from method 3.

In both spectral ranges (see Figure 4.6 and Figure 4.7), all the inks present in the historical manuscripts were classified as mixture of iron gall ink with sepia or lampblack. This is consistent with previous studies on these manuscripts [22], which identified an organic component mixed with iron gall ink. In addition, in the second minicube, some pure iron gall ink was found in the central area of the stroke, which could be attributed to irregularities in the ink composition. In the SWIR range (see Figure 4.7), the first and third minicubes were primarily classified as IGISE with SAM, while $SIDSAM$ and $JMSAM$ classified them as IGILA. This discrepancy

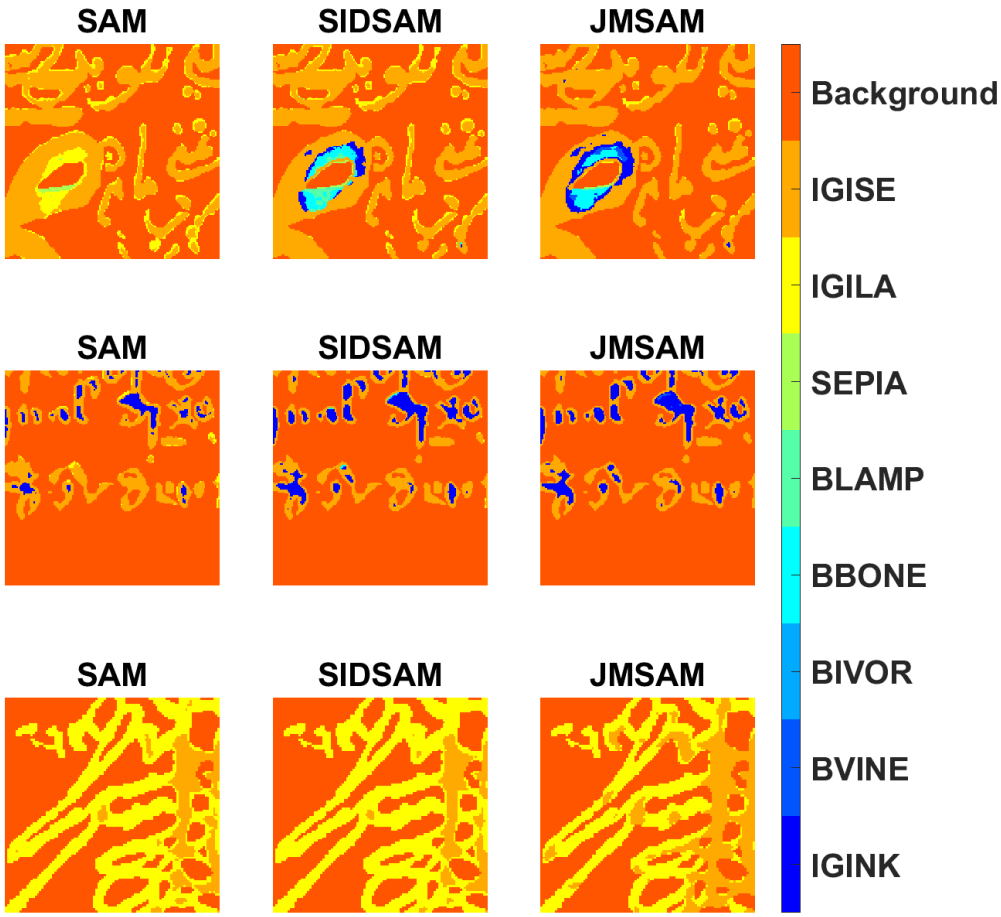

Figure 4.6: Classification maps of historical manuscripts using SAM, SIDSAM and JMSAM metrics in the VNIR range.

could be attributed to the fact that SAM does not consider magnitude shifts in the spectrum. In the VNIR range (see Figure 4.6), the results obtained with the three metrics were more comparable.

4.4 Discussion and conclusions

Three methods for selecting optimal spectral metrics for the purpose of classifying historical inks are used to evaluate six spectral similarity metrics. Hyperspectral images of mock-ups and real historical samples were acquired in the VNIR and

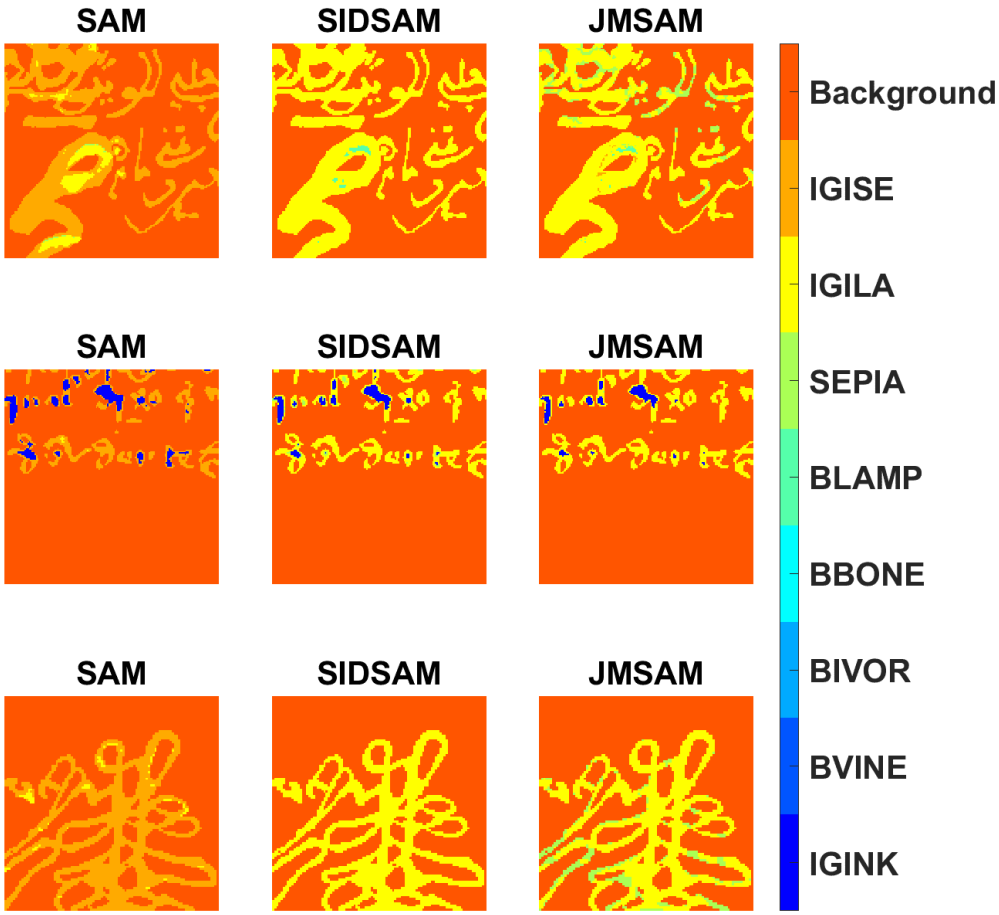

Figure 4.7: Classification maps of historical manuscripts using SAM, SIDSAM and JMSAM metrics in the SWIR range.

SWIR ranges. A comparative of the metrics was conducted using mock-ups to determine the most suitable ones. The selected metrics were applied to identify the ink present in samples of ancient manuscripts.

From the study of the mock-ups, considering the confidence thresholds selected, $SIDSAM$ outperformed the other metrics in both spectral ranges. Considering the classification performance, the metric that provided the highest accuracy and F1-score for both spectral ranges was SAM. These findings align with the results observed in a previous study [129], although other study reported a superior performance with SID in the task of classifying characters overlapped of two different

inks [131]. Finally, from the study of the PSD and $PWSD$, $SIDSAM$ and $JMSAM$ were selected as optimal metrics. $SIDSAM$, SAM, and $JMSAM$ were used to classify three historical documents of the XVth century, providing consistent results in the VNIR range. In conclusion, pooling the results of the three methods for evaluating spectral similarity metrics proved to be effective in correctly identifying inks in both mock-ups and historical samples, providing a simple and reliable approach without requiring a priori information about the document under study.

Similarity metrics, such as those explored in this chapter, often perform less effectively when the spectral differences between two samples are minimal. In chapter 5 the application of ML and DL methods for material identification is explored. In addition, it is worth noting that the spectral reflectance of inks from different classes may exhibit very similar spectra, and the same ink, due to aging, interaction with the support, or other alterations, can display different spectra across different documents. Some studies have attempted to model pigment or ink aging [10,32,164,272], but further efforts are required in this field. To address this challenge, it might be interesting to include spectra of real aged samples (historical documents) in the spectral library. Despite this, the results obtained in this study show promise considering the simple methods used for material identification.

CHAPTER 5

Ink classification in historical documents using HSI and machine learning methods

In this chapter, five traditional supervised ML algorithms and one DL-based were evaluated on HSI captures of both mock-up samples and historical documents to classify three categories of inks: pure metallo-gallate inks, carbon-containing inks, and non-carbon-containing inks. This work was conducted in collaboration with Marco Buzzelli, from the University of Milano-Bicocca, Italy, who implemented the DL-based model. It is the outcome of specific objective O4 of this PhD thesis.

This study has been published as:

A. B. López-Baldomero*, M. Buzzelli, F. Moronta-Montero, M. Á. Martínez-Domingo, and E. M. Valero, "Ink classification in historical documents using hyperspectral imaging and machine learning methods," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 335, p. 125916, 2025.

5.1 Introduction

As stated in chapter 4, the identification of inks is important for several reasons. Historically, different types of inks have been used across cultures and periods, as explained in chapter 2, section 2.1. The increasing attention to the material composition of manuscripts and ancient artifacts, in general, reflects a broader recognition of their importance. This PhD thesis is focused on material identification from HSI. However, unlike the techniques summarized in section 2.2, HSI spectra do not directly reflect the chemical composition of the materials in the document. This makes additional processing necessary to gain access to this information, generally by recurring to classification algorithms.

The high dimensionality of HSI data makes it particularly suitable for integra-

tion with ML models. The increasing use of ML techniques coupled to HSI data is shown in recent studies for different purposes [65,273–275]. In particular, supervised classification algorithms have been proposed for pigment classification, including SVM, PLS-DA, DT, and DL techniques, as further explained in subsubsection 2.5.1.2. However, most automatic algorithms developed for material classification in artworks have been trained and tested primarily on mock-ups, i.e., controlled samples created by researchers. As a result, these algorithms may not always perform accurately when applied to real works of art, which can exhibit significant variability due to aging, different states of conservation, and different compounds incorporated into the recipe for a given material class.

In the context of document analysis, several studies have focused on identifying contemporary inks using HSI in forensic analysis [221–223, 258, 259], as mentioned in chapter 4, as well as the analysis of pigments in illuminated manuscripts [134, 161, 276]. However, to our knowledge, the only study that addressed the classification of historical inks using spectral metrics and a reference library was the one presented in chapter 4 [224]. Thus, the classification of historical inks using non-invasive techniques has received limited attention, particularly when relying exclusively on HSI data.

To date, no study has investigated the automatic classification of historical inks by using ML methods and HSI data. Therefore, the objective of this study was to train and validate six state-of-the-art supervised ML models to automatically classify three types of inks: (1) pure metallo-gallate inks (MGP), (2) carbon-containing inks (CC), which include pure carbon-based inks like ivory black or bone black, as well as mixtures of carbon-based and metallo-gallate or sepia inks, and (3) non-carbon-containing inks (NCC), which can be pure sepia or a mixture of MGP and sepia. Throughout this study, the six algorithms will be divided into two groups:

five in the group of traditional techniques, including SVM, KNN, Linear Discriminant Analysis (LDA), Random Forest (RF), and PLS-DA, and one in the group of DL techniques. Given that all inks appear brownish or black in the visible spectrum, hyperspectral images in the VNIR, from 400 to 1000 nm, and SWIR, from 900 to 1700 nm, ranges were captured, and low-level fusion was performed in the spectral dimension to enhance classification accuracy. Both mock-ups and historical documents were included in the training and test sets.

In addition to the primary objective, PCA was used prior to classification for visualization of the separability of the classes and dimensionality reduction, comparing the classification accuracy and running time with and without PCA. Parameter optimization for the different models was also tackled, and for the traditional algorithms, a novel post-processing step to increase the local consistency of the results was developed. The full workflow for several traditional algorithms made their performance comparable (only slightly worse) to that of the DL model. The decision to use traditional or DL-based algorithms will then be made depending on the available resources for computation at a given site and the availability of training data. The results of our study show that it is possible to perform ink identification and mapping using only spectral information, thus adding to the evidence validating HSI as a key non-invasive technology in the domain of historical document characterization and preservation.

5.2 Materials and methods

5.2.1 Mock-up and historical samples

Both mock-up and historical samples were used to train and test different ML models.

The mock-up samples were extracted from the set of modern synthetic samples presented in subsubsection 3.2.1.1. These included metallo-gallate inks, sepia, and carbon-based inks, along with their mixtures, prepared according to different traditional recipes [236], and using materials of varying provenance to ensure a wide range of variability. Two writing supports were used: parchment and hand-crafted cotton-linen paper.

The historical documents were sourced from three different collections, as detailed in subsubsection 3.2.1.2, all preserved in the Provincial Historical Archive and the Royal Chancellery Archive of Granada, Spain.

Further details on the materials present in the mock-up and historical samples, along with the associated labeling for the classification task, are provided in Table 5.1. Additional information about the recipes used in the mock-up samples can be found in subsubsection 3.2.1.1. Note that the exact recipes for the historical documents are unknown.

5.2.2 Hyperspectral image acquisition

Hyperspectral images of mock-up and historical samples were acquired in the VNIR and SWIR spectral ranges as detailed in subsection 3.2.2, and reflectance hypercubes were generated as described in subsubsection 3.2.3.1.

Table 5.1: Details of the supports, ink types, and corresponding labels for the samples used in the study.

Set	Inks		Label	Supports
Mock-up samples	Metallo-gallate - ferrous sulfate in different proportions		MGP	Parchment / Cotton-linen
	Metallo-gallate - FeSO4 + CuSO4		MGP	
	Metallo-gallate - FeSO4 + ZnSO4		MGP	
	Metallo-gallate - FeSO4 + pomegranate juice		MGP	
	Metallo-gallate – FeSO4 + pomegranate juice + myrtle leaves infusion		MGP	
	Metallo-gallate - FeSO4 + earth pigment*		MGP	
	Atramentum*		MGP	
	Ivory black*		CC	
	Bone black*		CC	
	Lamp black*		CC	
	Grape seed black*		CC	
	Cherry black*		CC	
	Bistre*		CC	
	Metallo-gallate - FeSO4 + lamp black* in different proportions		CC	
	Metallo-gallate - FeSO4 + bone black* in different proportions		CC	
	Lamp black* + earth pigment*		CC	
	Lamp black* + sepia in different proportions		CC	
	Bone black* + sepia in different proportions		CC	
	Sepia (from the ink sac of the animal) in different proportions		NCC	
	Sepia in powdered form*		NCC	
	Metallo-gallate - FeSO4 + sepia in different proportions		NCC	
Historical documents	1st set (notarial documents)	Metallo-gallate	MGP	Linen
		Metallo-gallate + earth	MGP	
		Carbon-based ink	CC	
		Carbon-based ink + earth	CC	
	2nd set (family tree book)	Metallo-gallate + sepia	NCC	Cotton-linen
		Carbon-based ink	CC	
	3rd set (lawsuits of nobility)	Metallo-gallate	MGP	Parchment

*From Kremer Pigmente GmbH.

5.2.3 Data pre-processing - Registration, sample extraction, Ground Truth images, and data fusion

Data curation was performed on the reflectance hypercubes, which included spatial registration of spectral data from both the VNIR and SWIR HSI cameras, extraction of representative areas containing support and one or two different types of ink (referred to as *minicubes*), and the creation of GT images with per-pixel material

identification for each minicube, as described in subsection 3.2.3.

An example of spatial registration is shown in the first row of Figure 5.1. Figure 5.1 d) and e) display the false-color images of the minicube in the VNIR (R = 645 nm, G = 565 nm, and B = 440 nm) and SWIR (R = 1600 nm, G = 1200 nm, and B = 1000 nm) ranges. Figure 5.1 f) shows the GT image for the minicube, with the presence of two different inks: metallo-gallate ink mixed with sepia (orange) and carbon-based ink (yellow). The background pixels are represented in middle gray (with RGB values set to 128).

After extracting the minicubes and building the GT images, the spectral data in both ranges were fused. The fusion process is performed by integrating different data sources to produce more useful and accurate information than any individual data source [277]. In our case, a low-level fusion is performed, where the VNIR and SWIR spectra are concatenated using the bands 400-950 and 955-1700 nm, respectively, without further pre-processing [278]. This resulted in fused spectra with 261 bands. Although PCA could have been applied before data fusion for a mid-level approach, this would carry the risk of discarding potentially important information during dimensionality reduction [278], which could be critical for distinguishing ink classes with overlapping spectral properties. Figure 5.1 g) shows the data fusion result for a single pixel of metallo-gallate ink mixed with sepia (NCC group), illustrating the contribution of each range.

5.2.4 Training and testing sets

The training and test sets were extracted from a full dataset comprising 44 registered pairs of VNIR and SWIR documents described in subsubsection 3.2.1.2. From these, 145 minicube pairs were extracted.

For our experiments, the documents were partitioned into two sets, so that the

corresponding minicubes cover 75% of the total for training and 25% for test. Partitioning at document level allows different minicubes extracted from the same document to fall into the same subset. Such precaution prevents the introduction of bias in the training-test split, avoiding test minicubes from having training counterparts coming from the same document. In fact, this would create an unrealistic evaluation scenario, where test performance does not reflect real-world performance.

For the training and test sets of traditional models, only ink spectra (according to the GT images) were selected. This includes pixels from square regions with higher ink deposition in mock-ups, as well as traces with variable amounts of ink, allowing the models to account for these variations during classification, ensuring robustness in the results. This does not exclude the possibility that a particularly dark support pixel may have been included in the training data incorrectly, but this will not happen often because the GTs were carefully revised.

Following this procedure, we obtained the data distribution shown in Table 5.2. Due to computational workload limitations, the number of pixels in the training set was reduced by randomly subsampling at a 1:5 ratio after data shuffling. This reduction was necessary for the parameter optimization of traditional algorithms, as detailed in subsection 5.2.7. After the subsampling step, the number of pixels per class was reduced without altering the class imbalance, as shown in Table 5.2.

Hence, the three classes are not balanced, with 44% of the spectra (pixels) in the CC class, 31% in the MGP class, and 25% in the NCC class. It is worth noting that the sum of the samples containing each class exceeds the total number of samples in the table, as some samples include two types of ink instead of just one. For traditional algorithms, class imbalance compensation by reducing the majority classes did not improve the classification results in preliminary experiments.

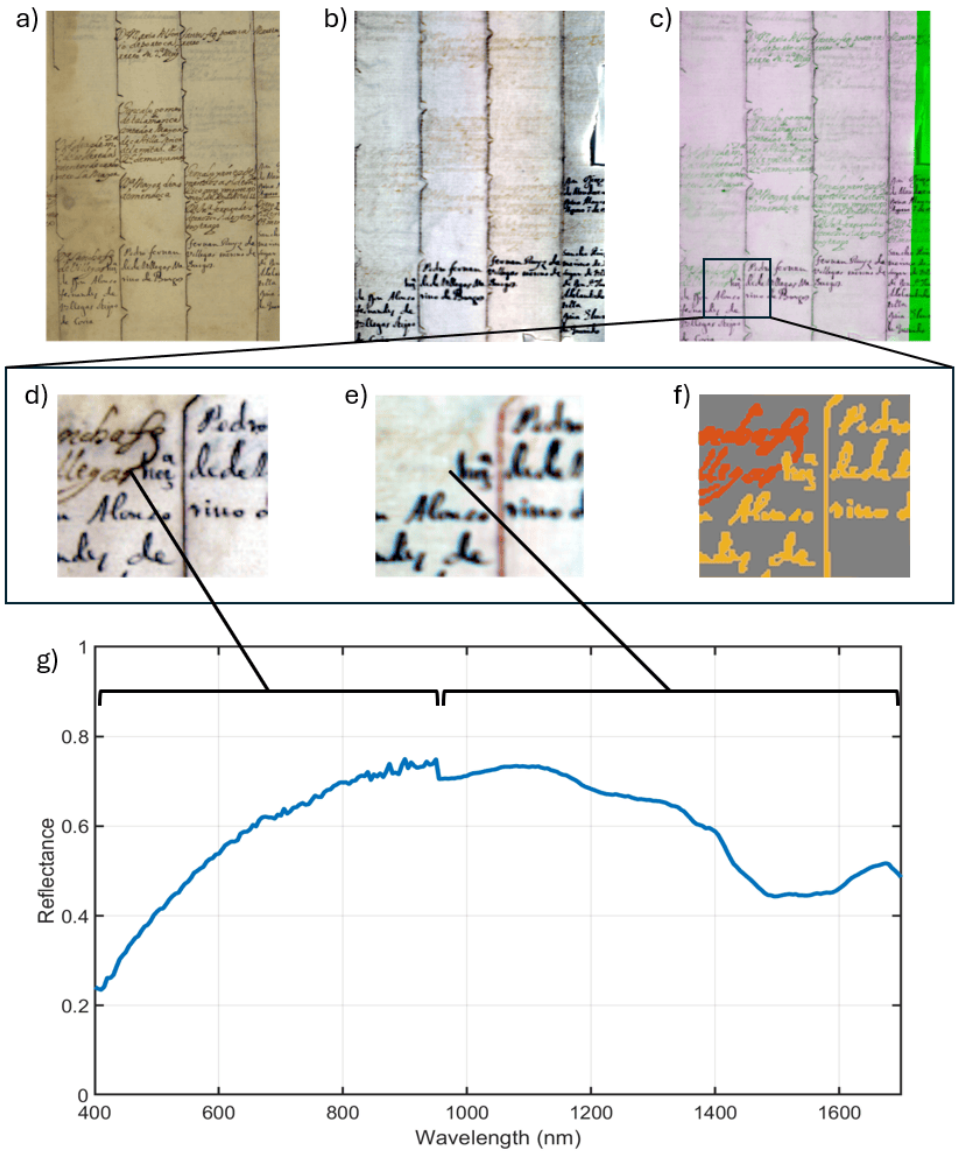

Figure 5.1: Registration process and outcome. a) and b) False color images in the VNIR (R = 645 nm, G = 565 nm, B = 440 nm) and SWIR (R = 1600 nm, G = 1200 nm, B = 1000 nm) ranges, respectively. c) Overlay of both images, with green indicating areas belonging only to the SWIR capture. d) and e) False color images of registered VNIR and SWIR minicubes, respectively. f) Ground Truth. g) Spectrum of an ink pixel after data fusion.

Table 5.2: Training and test data distribution for the three ink classes: pure metallogallate inks (MGP), carbon-containing inks (CC), and non-carbon-containing inks (NCC).

	Train			**Test**		**Total**	
Class	**Samples**	**Pixels**	**Pixels subsam.**	**Samples**	**Pixels**	**Samples**	**N. pixels**
MGP	49	232534	46510	14	71349	63	303883
CC	45	330236	66114	17	210016	62	540252
NCC	28	188497	37833	7	24890	35	213387
Total	109	751267	150457	36	306255	145	1057522

5.2.5 Principal Component Analysis (PCA)

In this chapter, PCA is employed for both dimensionality reduction and visualization purposes. First, it is used to reduce the amount of data introduced to the classifiers to improve efficiency, comparing the accuracy and running time for the training phases of the models with and without PCA, to determine whether PCA is useful for reducing training time without compromising accuracy. To do that, the optimal number of PCs was selected, and their projection coefficients were used to train the classifiers, as explained in subsection 5.2.7.

Additionally, PCA is applied to the hyperspectral data of the full training set for visualization, projecting the first principal components of each spectrum onto a 2D graph. This approach facilitates a quick assessment of the separability of the data, allowing for the visualization of the three separate classes and helping to identify whether distinct clusters can be found, as explained in section 5.3.

5.2.6 Classification models

All models used in this study are supervised classification techniques, so prior information about the data is required. These algorithms automatically identify spectral signatures corresponding to various types of inks, facilitating the classifi-

cation of unknown samples through the use of a reference or training dataset. As explained before, two groups of algorithms are considered: the traditional and the DL-based. For the implementation of the traditional models, MATLAB software (release R2023a, The MathWorks, Inc., Natick, MA, USA) was used. For the implementation of the DL-based model, Python 3.10.12 was used with the PyTorch deep learning framework at version 1.11.0.

Table 5.3 provides a summary of the five traditional algorithms (SVM, KNN, LDA, RF, and PLS-DA) and the DL-based model, highlighting their fundamentals, advantages, limitations, and the hyperparameters used for each. For details on hyperparameter optimization, please refer to subsection 5.2.7.

Table 5.3: Summary of classification models used in the study: fundamentals, advantages, limitations, and hyperparameters.

Model	Fundamentals	Advantages	Disadvantages	Hyperparameters	References
Support Vector Machines (SVM)	Finds the optimal hyperplane that separates classes with the maximum margin. Handles both linear and nonlinear problems using kernel functions.	- Robust theoretical framework. - Efficient for high-dimensional data. - Strong generalization. - Minimizes the risk of overfitting.	- Requires careful parameter tuning. - Computationally expensive for large datasets. - Originally designed for binary classification.	- Kernel function = Gaussian - Box constraint level = 10 - Kernel scale = Automatic - Multiclass coding = One-vs-One - Standardize data = Yes	[157] [65,275]
K-Nearest Neighbors (KNN)	Classifies based on the majority vote of K nearest neighbors; no training phase required.	- Transparent and interpretable. - Well-suited for multi-class problems. - No assumptions about data distribution.	- Sensitive to noise and spectral variability. - High computational cost for large datasets. - Performance highly dependent on K value and distance metric.	- Number of neighbors = 1 - Distance metric = Cosine - Distance weight = Equal - Standardize data = Yes	[275] [279,280]
Linear Discriminant Analysis (LDA)	Finds a linear combination of features that best separates two or more classes by maximizing the ratio of between-class variance to within-class variance. Projects the data onto a lower-dimensional space.	- Computationally efficient for large datasets. - Easy to implement and interpret. - Robust to overfitting. - Few hyperparameters. - Well-suited for multi-class problems.	- Assumes linear separability, normal distribution, and equal covariance matrices for all classes. - Performance degrades with high-dimensional or noisy data.	- Discriminant type = Linear - Covariance structure = Full	[274,281]

Table 5.3: (continued) Summary of classification models used in the study: fundamentals, advantages, limitations, and hyperparameters.

Model	Fundamentals	Advantages	Disadvantages	Hyperparameters	Ref.
Random Forest (RF)	An ensemble of decision trees trained on random subsets of data and features. Uses majority voting for classification.	- Robust to noise and overfitting. - Handles high-dimensional data effectively. - No assumptions about data distribution. - Handles collinearity well.	- Sensitive to hyperparameter choices. - Increased complexity with many trees. - Less interpretable compared to simpler models. - High computational cost.	- Ensemble method = Bag - Learner type = Decision tree - Maximum number of splits = 751266 - Number of learners = 30 - Number of predictors to sample = Select All	[275, 280]
Partial Least Squares Discriminant Analysis (PLS-DA)	Combines PLS regression and LDA to enhance class separation by creating discriminant functions from input variables that provide better separation than individual variables alone.	- Handles collinearity. - Effective for small sample sizes. - Provides robust class separation.	- Struggles with noisy or highly complex data. - Requires careful pre-processing.	- Preprocessing = Autoscale - Number of latent variables = 3 - Orthogonalize = Off - Algorithm = SIMPLS - Sample weighting = None	[43, 273]
Deep Learning (DL)	Uses the DeepLabV3 semantic segmentation model for pixel-wise classification, employing dilated convolutions to efficiently integrate multiscale information.	- Excellent for high-dimensional data. - Models complex relationships and spectral variability. - Benefits from transfer learning.	- High sensitivity to noise without proper training. - Requires large computational resources and datasets. - Increased risk of overfitting if not carefully tuned.	- Epochs = 25 (total of 3300 iterations) - Loss function = categorical cross-entropy - Adam optimizer - Learning rate = 0.0001.	[282, 283]

PLS-DA was implemented using the PLS_Toolbox (Eigenvector Research, Wenatchee, US). For the DL-based algorithm, neural parameters were initialized through pretraining on a subset of the Microsoft COCO dataset [284]. Although this dataset depicts a different type of visual content, scientific literature suggests that exposing the model to diverse visual data can improve training speed and reduce the amount of required training data [283]. To adapt the model architecture to our problem, we replaced the first convolutional layer, which was originally designed to process 3-channel RGB images, with a new convolutional layer capable of processing 261 channels (111 VNIR + 150 SWIR channels).

5.2.7 Optimization and post-processing for traditional algorithms

A k-fold cross-validation method with k = 5 was employed to optimize some of the traditional model parameters using a subsampled training set (see Table 5.2). This technique divides the dataset into training and testing subsets as explained in chapter 3, subsection 3.4.1.

For KNN optimization, six distance metrics were evaluated: cityblock, Chebychev, correlation, cosine, Euclidean, and Minkowski. First, the optimal distance metric was determined using a number of neighbors K = 1. After that, a different number of neighbors were tested with the optimal distance metric: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, and 100. A low K value can lead to overfitting, where the model memorizes the training data too closely and performs poorly on new, unseen data. On the other hand, a high K value can result in underfitting, where the model fails to capture the underlying patterns in the data adequately [280]. The performance metrics data obtained with the subsampled training set in cross-validation (k = 5) can be found in Table 5.4. From these results, the cosine distance and K = 1 were selected as the final hyperparameters. Although comparable performance was achieved with K values

Table 5.4: Micro-averaged accuracy of KNN with different distance metrics and numbers of neighbors.

Model	Distance metric	Neighbors	Micro-accuracy
KNN	cityblock	1	0.9817
KNN	chebychev	1	0.9764
KNN	correlation	1	0.9794
KNN	cosine	1	0.9845
KNN	euclidean	1	0.9826
KNN	minkowski	1	0.9826
KNN	cosine	2	0.9817
KNN	cosine	3	0.9832
KNN	cosine	4	0.9825
KNN	cosine	5	0.9823
KNN	cosine	6	0.9816
KNN	cosine	7	0.9813
KNN	cosine	8	0.9808
KNN	cosine	9	0.9802
KNN	cosine	10	0.9798
KNN	cosine	20	0.9759
KNN	cosine	50	0.9685
KNN	cosine	100	0.9613

of 3, 4, and 5 using cosine distance, a significant increase in K led to a noticeable decline in performance.

For the SVM model, the box constraint was optimized using the following values: 0.8, 1, 1.2, 1.4, 1.6, 2, 4, 10, 30, 100, 200, and 300. In this case, to select the best value, both micro-accuracy and training time were considered, an original parameterization approach proposed in this thesis. Increasing the box constraint results in the SVM classifier assigning fewer support vectors, which leads to stricter data separation, but also to longer training times [285]. After evaluating the training time in cross-validation (k = 5) and micro-averaged accuracy (see Figure 5.2), a box constraint of 10 was selected.

After obtaining the classification maps as explained in subsection 5.2.8, a post-processing cleaning procedure was applied. Considering the assumption that a con-

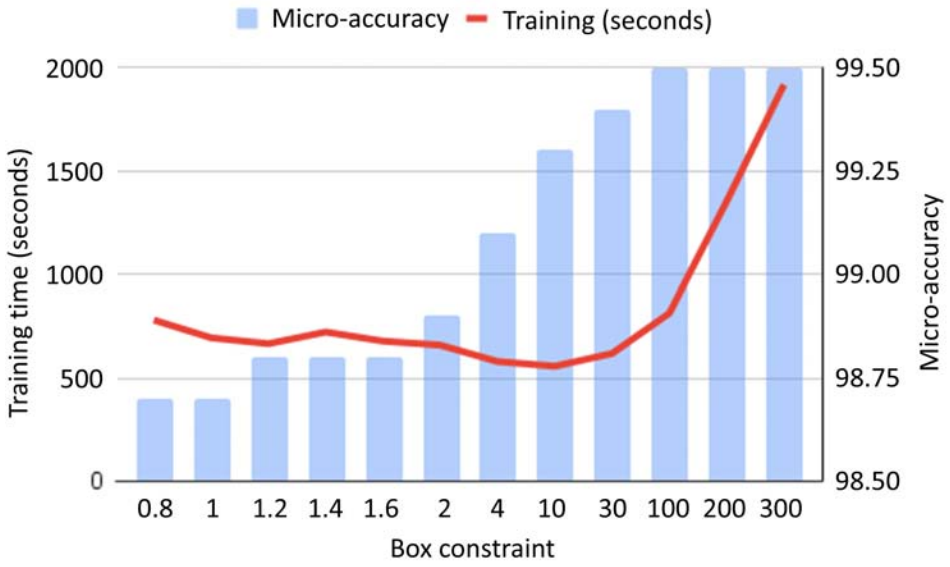

Figure 5.2: Training time in seconds (red line) and micro-averaged accuracy (bar charts) after cross-validation (k = 5) for different values of the box constraint in the SVM model.

tinuous stroke is composed of the same ink type, each pixel in the classification map was reassigned to the most prevalent class within its surrounding neighborhood. Neighborhood size was defined as 5% of the smallest dimension of the minicube. The cleaning process was repeated over 10 iterations. These parameters were selected based on preliminary tests that indicated optimal performance with minimum computational time. The post-processing step is one of the contributions of this study, and its impact on the performance of traditional models is described in section 5.3.

5.2.8 Performance evaluation

To evaluate the performance on the test set, the confusion matrix was first employed to compute pixel-level performance metrics. The number of TP, FP, TN, and FN were used to calculate micro-averaged accuracy and macro-averaged precision, recall, accuracy, and F1-score, as described in subsubsection 2.5.1.2.

Additionally, classification maps were generated for each minicube based on the prediction results, with each class represented by a distinct color group to facilitate the quick identification of misclassifications. The color code used is: purple for MGP, yellow for CC, and orange for NCC. These maps were then visually inspected to assess the consistency of the classifications.

5.2.9 Case study: binarization and classification of inks in two full historical documents

After evaluating the different models, a practical application is presented to demonstrate a complete classification process of a historical document using the best-performing models (either traditional or DL-based).

For this purpose, two documents, one from the Royal Chancellery Archive and another from the Provincial Historical Archive of Granada, were selected. The first document is a page from a family tree book dating from between the 16th and 17th centuries. This page is entirely handwritten, and the hands of two different people can be identified in it, each using a particular ink. In this case, the reason for the existence of two authors is unknown. In previous analyses conducted by the conservators in charge, it has been verified the presence of a CC ink and another ink consisting of a mixture of MGP and sepia (NCC). Two minicubes extracted from this document were used as part of the training set samples.

The second document is an Arabic notarial manuscript dated to 1499, detailing a certificate of ownership for irrigated land in the Hotallar village. Such documents are particularly valuable for classification studies, as they typically contain the handwriting of two individuals: a notary who writes the document and a judge who validates it, adding a few words to indicate his agreement, each using different inks.

In this manuscript, the text and marginal note are written with a MGP-based ink (MGP), and the judge's validation with a pure carbon-based ink (CC). No minicubes from this document were included in the training set.

For these two documents, an additional pre-processing step was required: the binarization of the document. This step consists of separating the background (support) and foreground (inks), so that we can use only the ink pixels as input to the classifiers. The spectral band with the highest contrast was selected using the SNR metric as explained in subsubsection 3.2.3.4, and then Bradley's Local Image Thresholding algorithm [252] was applied. This method chooses a threshold T for each pixel based on its surroundings:

$$T = \mu \cdot \left(1 - \frac{t}{100}\right) \tag{5.2.1}$$

where μ represents the local mean intensity within the chosen window, and t is the percentage of intensity values to be considered as foreground. In our case, we used a window size of $\frac{1}{3}$ of the image height times $\frac{1}{3}$ of the image width, and t is set to 10. We have selected this algorithm and parameters since they obtained the best results in a previous study with similar documents [231].

After classifying the pixels that Bradley's method selected as ink, we performed the same cleaning post-processing as described in previous sections for the traditional algorithms.

To facilitate the evaluation of the classification maps, a GT was manually created using GIMP 2.10.38 software based on the binarized images. It is important to note that these GTs are intended to provide a general overview of the inks in the areas rather than a precise pixelwise identification as performed for the minicubes, as minor manual errors may be present. Therefore, our focus was on verifying that

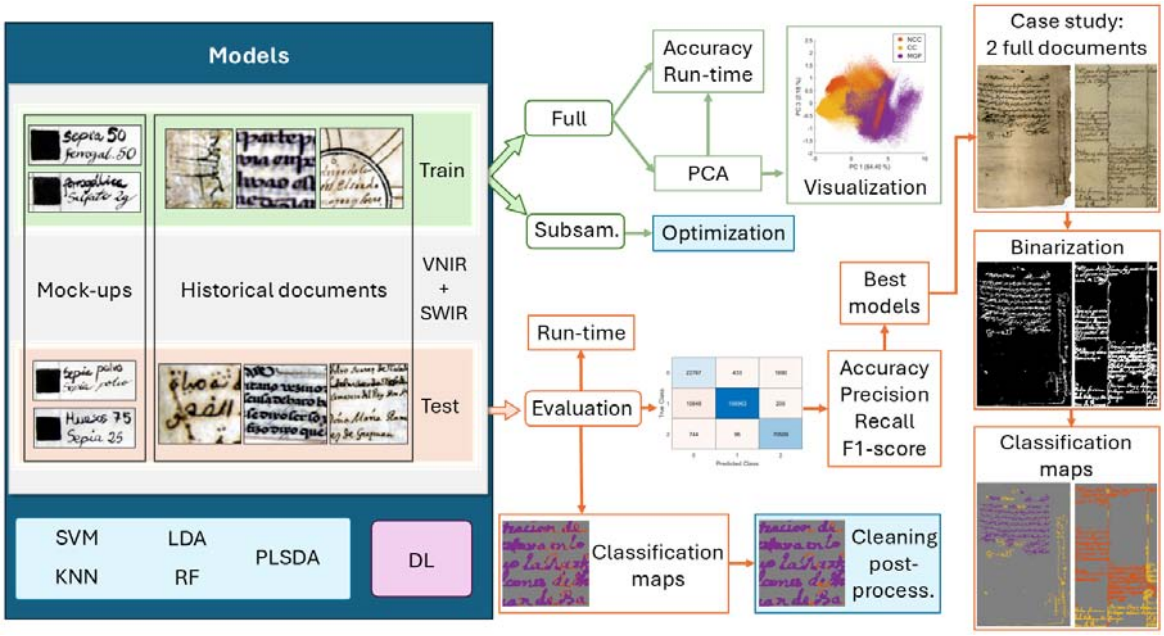

Figure 5.3: Workflow outlining the steps followed during the different phases of this study.

the number of ink types and their relative spatial positions were consistent with the findings from previous analyses of the documents.

Figure 5.3 presents the workflow of the methodology followed in this study, providing a visual representation of the procedures outlined in the preceding subsections. Optimization and cleaning post-processing are shaded in blue as they were only performed for traditional algorithms.

5.3 Results and discussion

5.3.1 Average spectral reflectance of inks

In Figure 5.4, the average spectral reflectance and standard deviation of the three ink classes from 400 to 1700 nm are presented, along with the average reflectance of two of the writing supports, parchment and cotton-linen paper. The ink spectra

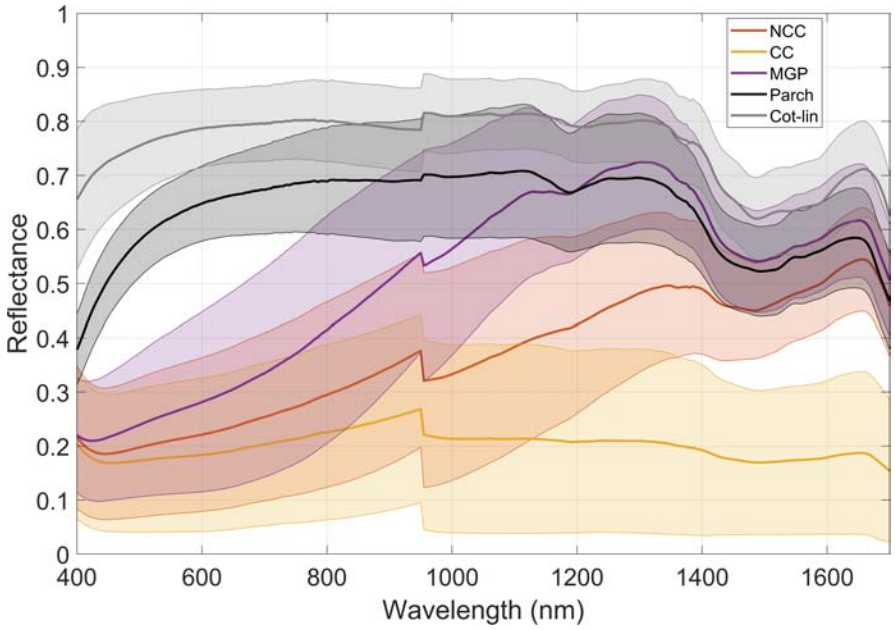

Figure 5.4: Average spectral reflectance and standard deviation for the three ink classes in the full training set and two writing supports: non-carbon-containing inks (NCC), carbon-containing inks (CC), pure metallo-gallate inks (MGP), parchment (Parch), and cotton-linen paper (Cot-lin).

were extracted from the pixels marked as ink in the GT images of the full training set. An offset can be observed in the 950 to 955 nm range, where data fusion occurred. This is common when different sensors are used for capture and is caused by several factors, including differences in spectral bandwidths, low signal-to-noise ratios, and misalignments in the image setup, which slightly affect the BRDF [255]. In the visible range, the reflectance patterns of the three inks are similar, showing very low values and a flat shape (with a trend toward reddish color for the MGP ink). However, as we explore into the near-infrared region, the reflectance of MGP ink diverges, increasing notably as seen in previous studies [286,287]. This divergence is particularly evident starting at approximately 1300 nm, where pure MGP ink becomes nearly transparent (i.e. it lets the infrared radiation pass through almost completely, and what one sees in this spectral range of the reflectance curve is the reflectance of the support). This

near transparency distinguishes MGP ink from other inks, including CC ink and sepia, as well as mixtures of inks with or without carbon content. Specifically, CC inks absorb a significant amount of infrared radiation. However, pure sepia and the mixtures of sepia and MGP ink (included in the NCC class) allow slightly more infrared radiation to pass through but do not reach the near-total transparency seen in the MGP class spectra.

In the spectral range from 400 to 1700 nm, the infrared region is particularly interesting as molecular overtone and combination vibrations can be studied from it. Specific absorption bands within this range are associated with distinct chemical bonds: the 1460–1570 nm range corresponds to N-H bond absorptions, the 1100–1400 nm and approximately 1700 nm bands are attributed to C-H bond absorptions, and the 1450 nm band is linked to O-H bond absorptions [288]. Considering the shape of the reflectance curves in Figure 5.4, MGP and NCC inks have two peaks in the infrared range: one around 1300-1400 nm, and the other at 1650 nm approximately. In CC inks, the latter peak can also be seen, but much less pronounced. Both peaks could be reasonably assumed to correspond to C-H bond absorptions. However, in this spectral range the absorption bands are weaker and more complex than those in the mid-infrared region, so the application of chemometric techniques is highly suitable to achieve a higher level of confidence in the classification of ink spectra.

5.3.2 PCA for visualization

In Figure 5.5, score plots of principal components (PCs) 1, 2 and 3 are presented. Three PCs were selected by analyzing the Variance Accounted For (VAF) curve and identifying the inflection point at which the curve flattens out. As seen in Figure 5.5, 84.4% of the total variance is explained by PC1, 12.1% variance is explained by PC2, and 2.2% is explained by PC3, achieving a total of 98.6% VAF with just three compo-

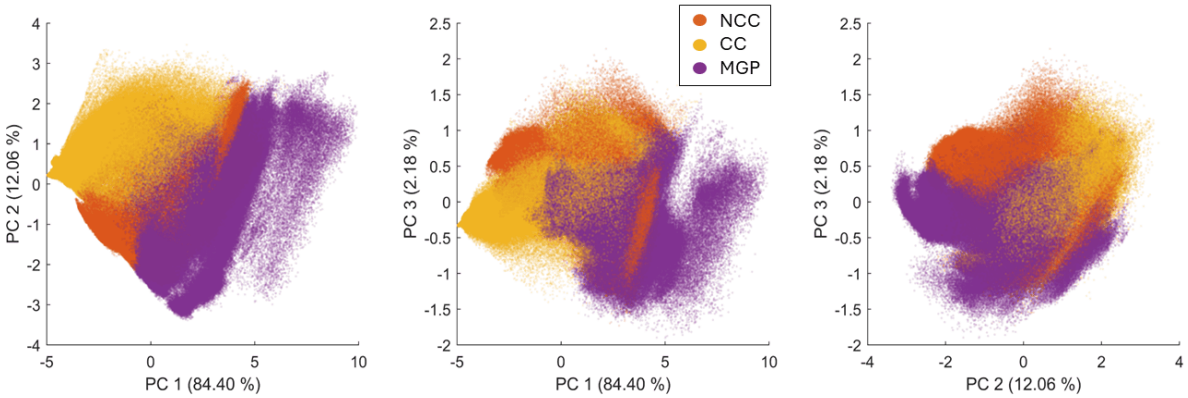

Figure 5.5: Score plots of principal components (PC) 1, 2 and 3 for the three different classes used in the study: non-carbon-containing inks (NCC), carbon-containing inks (CC), pure metallo-gallate inks (MGP).

nents. In this Figure, the point cloud for MGP inks, represented in purple, seems to form a separable cluster from that of CC inks, represented in yellow. However, the point cloud for the NCC class is located between the two previous groups. It should be noted that the CC group contains pure carbon-based inks as well as mixtures with metallo-gallate and sepia inks. Similarly, the NCC group includes pure sepia ink and its mixture with MGP ink. This may be the reason why no clear clustering pattern is observed among the three groups, from which it is concluded that the data are not clearly separable in the PCA components space.

5.3.3 Classification maps and performance metrics

In Table 5.5, the performance metrics for all traditional classification models evaluated on the test set before applying the cleaning post-processing step are presented. SVM provided the highest performance across all the metrics studied, including both micro- and macro-averaged results, achieving over 95% in micro and macro-averaged accuracy and recall. This superior performance could be attributed to its efficiency in handling high-dimensional data, its ability to model non-linear rela-

Table 5.5: Performance metrics in the test set for all traditional models before cleaning post-processing. The color shades represent a gradient from best (dark green) to worst (dark red) for each column.

Model	Micro-accuracy	Macro-accuracy	Macro-precision	Macro-recall	Macro-F1
SVM	96.06	95.38	89.17	95.38	91.79
KNN	93.22	93.39	84.09	93.39	87.43
LDA	95.27	93.02	87.90	93.02	89.73
RF	94.77	94.33	86.72	94.33	89.70
PLSDA	86.82	78.33	72.42	78.33	74.36

Table 5.6: Performance metrics in the test set for all models after cleaning post-processing. The color shades represent a gradient from best (dark green) to worst (dark red) for each column.

Model	Micro-accuracy	Macro-accuracy	Macro-precision	Macro-recall	Macro-F1
SVM	98.35	97.29	95.08	97.29	96.15
KNN	95.23	95.79	87.36	95.79	90.63
LDA	97.42	95.44	92.40	95.44	93.71
RF	97.25	97.13	91.78	97.13	94.14
PLSDA	89.80	81.56	76.09	81.56	78.23
DL	99.20	99.13	97.40	99.13	98.22

tionships, and its robustness to overfitting (see Table 5.3). In contrast, PLS-DA is the model providing the lowest values for all the metrics. This suggests that a nonlinear model, like SVM, or a more flexible model, such as RF, may be more suitable for this problem. However, it should be noted that the values of all the metrics are above 72%, reaching almost 87% in micro-averaged accuracy. This means that even the worst of the models tested here provides what could be considered good results in this classification task (in comparison to random class assignment, which would yield only about 33% accuracy). All other models provide micro- and macro-averaged accuracy and recall above 93%.

Table 5.6 presents the performance metrics for the six models evaluated on the test set, after applying the cleaning post-processing step for the five traditional models. When compared to Table 5.5, a 2 to 3% improvement in micro and macro-averaged accuracy and recall can be observed after applying post-processing. In

addition, between 3 and 5% improvement in macro-precision and F1 is achieved. These results indicate that cleaning post-processing is beneficial for the problem we are addressing in the context of traditional models, since in the same stroke (and, therefore, in contiguous pixels in the hyperspectral image) it is not normal to find different types of inks. The post-processing helps improving the results obtained for all the performance metrics studied.

The DL model is included in the post-processing set of results since, by design, it exploits pixel neighborhood information to inform the final class prediction. It outperforms all traditional models with the post-processing step included, having both micro-accuracy and macro-recall above 99%. However, it requires specialized hardware in order to efficiently complete the training and inference phases.

Hyperspectral analysis combined with Least Squares SVM classification has been used for ink analysis and pen verification in handwritten documents, achieving an 87.5% accuracy in discriminating between 25 different pens with modern inks [223]. Another study analyzed 70 hyperspectral images of handwritten notes by 7 subjects, comparing 5 varieties of blue ink and 5 varieties of black ink, with a focus on ink mismatch detection [221]. However, these studies are not directly comparable to the present work, as they involve modern inks and are designed for forensic purposes.

A comparative analysis of performance by class based on the confusion matrices (see Figure 5.6) reveals that RF, SVM, and KNN provide the lowest macro-average performance metrics for the NCC class. These models provide also the highest accuracy and recall for the MGP class, and the highest precision and F1 for the CC class. The higher precision but lower recall for the CC class indicates that CC inks are more likely to be classified as MGP and NCC, than NCC are likely to be classified as CC. This misclassification can be attributed to the presence of mixtures of pure carbon with MGP or pure carbon with sepia (which is included in the NCC class). In con-

trast, both LDA and PLS-DA demonstrate the highest macro-average metrics for the CC class. PLS-DA, in particular, performs poorly for the NCC class, with a precision of 42.7% and an F1 score of 50%. This may be due to the inherent assumptions of the PLS-DA algorithm, which make it less effective at handling the complexity of this class, as PLS-DA is not well-suited for highly complex datasets (see Table 5.3 fifth row). In contrast, the metrics for the remaining two classes exceed 85%. This performance degradation may be due to class imbalance, as the reduction occurs in the least-represented class. Another explanation is that the spectrum of the NCC class is right between those of MGP and CC (see Figure 5.4), leading to increased misclassification between NCC and these two classes, compared to direct misclassification between CC and MGP.

The DL model performs favorably throughout all classes. The lowest recall is at 98.1% for MGP, which tends to be misclassified as NCC, in turn lowering its precision to 92.68%. Of all models tested, the DL achieves the best results for the NCC and CC classes, but not for the MGP class.

In general, for traditional models, NCC pixels tend to be misclassified as MGP and vice versa, while misclassification as CC is less frequent for both classes. However, for all the models when CC pixels are misclassified, they are more likely to be assigned to the NCC class than to MGP. This makes sense, as CC group includes some mixtures of sepia and carbon-based ink, but MGP only includes pure metallogallate inks with no mixtures with sepia. In addition, we have seen in Figure 5.4 that spectrally, NCC is more similar to CC inks than MGP.

Comparing different performance metrics, the lowest values were always obtained for macro-precision, due to the increased number of false positives for the NCC class. The highest values were obtained for the micro-averaged accuracy metric, which makes sense as MGP and CC are the most represented classes (as seen in

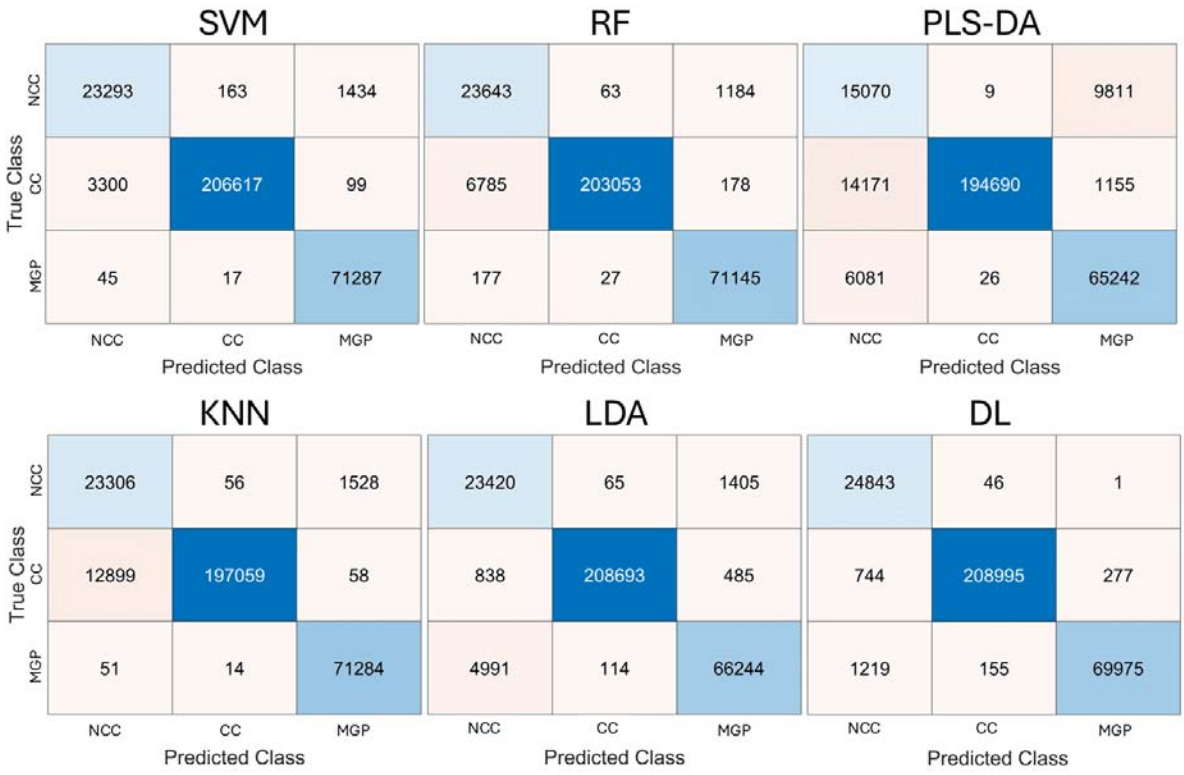

Figure 5.6: Confusion matrices of all classification models after cleaning post-processing for the test set. Darker blue in the diagonal cells indicates a higher number of TP spectra.

Table 5.2), and provide a high accuracy value.

In chapter 4, the classification of historical inks was performed using a library of reference spectra and different spectral metrics for pixel-by-pixel classification. Two spectral ranges, VNIR and SWIR, were studied separately, achieving a maximum F1-score of 58.1% for the VNIR range and 44.3% for the SWIR range using the SAM metric. However, the ink classes analyzed in the previous chapter differ from those in the present study, as carbon-based inks from different sources (e.g., vine black, ivory black, bone black, lamp black) were considered separately [224]. One key aspect contributing to the success of the classifiers used in this chapter has been the simplification of the number of classes, as the models still face difficulties

Table 5.7: Training run-time and micro-averaged accuracy comparison of different classifiers on the full dataset, with and without PCA.

	Full set (no PCA)		**Full set (PCA)**	
	Micro-accuracy	**Training time***	**Micro-accuracy**	**Training time***
SVM	99.51	7 h 37 min 56 s	92.33	4 h 19 min 24 s
KNN	99.17	25 h 7 min 19 s	85.61	48 min 55 s
LDA	96.73	6 h 52 min 21 s	86.49	12 s
RF	99.03	4 h 42 s	92.68	25 min 38 s
PLSDA	91.34	2 min 17 s	86.26	15 s
DL	99.85	1 h 4 min 50 s	99.32	1 h 1 min 58 s

*Computational environment used for experiments available in Table 5.8.

in intra-class separations (e.g., mixtures of sepia and MGP with CC are particularly challenging to distinguish).

In Table 5.7, training run-time and micro-averaged accuracy for the full training set with and without applying PCA are presented. By applying PCA, the number of features was reduced from 261 to 3, decreasing the training set to 1.15% of the original size. This resulted in a reduction of training run-time by 97-98% for most models, with the exception of PLS-DA, which showed a reduction of 89%, SVM that presented a reduction of 50%, and DL which introduced a negligible reduction of training run-time. However, applying PCA results in a 5-13% decrease in micro-averaged accuracy for traditional models, while the DL model is much more robust to the dimensionality reduction, likely due to the learned ability to extract complex relationships among the input features, and to the possibility of accessing neighborhood data directly during the inference phase. This trade-off between reduced training time and decreased accuracy should be considered when using traditional models for ink classification: if minimizing training time is prioritized over accuracy, then PCA can be applied. On the other hand, the reduction in time is much less significant for the DL model, which might not make the use of dimensionality

Table 5.8: Computational environments used for traditional and DL-based algorithms.

	Traditional	**DL-based**
Processor (CPU)	Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz (12 CPUs)	Intel(R) Core(TM) i7-7700 CPU @ 3.60GHz (8 CPUs)
Memory (RAM)	16 GB	32 GB
Storage	512 GB NTFS SSD	3 TB ext4 SSD
Operating System	Windows 11 Pro, v. 23H2, 64-bit	Ubuntu 22.04.3 LTS, 64-bit
Graphics Card (GPU)		NVIDIA Titan X, 12 GB

reduction worthwhile.

The run-time values of traditional models and DL cannot be directly compared because the DL model was run on a GPU, while the traditional models used CPU resources. Besides, the computers used for the two kinds of models were different (see Table 5.8 for details). However, comparing the training run-time between traditional models is possible: the fastest model to train was PLS-DA, followed by RF. The slowest model was KNN. However, it should also be clarified that run-time in KNN is related to the time required by the program to store the training dataset in the model, since this model does not have a training step as such (see Table 5.3). The full training dataset with all 261 bands was selected as the preferred set-up, as it provided the highest accuracy.

Additional insights into the model performances can be gathered from the classification maps. In Figure 5.7, some of these maps for selected mock-ups and historical samples are presented for SVM, PLS-DA, and DL models.

For the mock-up samples (first and second rows in Figure 5.7), some problems persist for the SVM model when differentiating between pure sepia ink and MGP (1st row). For the PLS-DA model, this problem was mostly solved after applying cleaning post-processing (first row, columns three and four). In the case of CC ink (2nd row), SVM misclassifies some pixels in the strokes as NCC, while PLS-DA

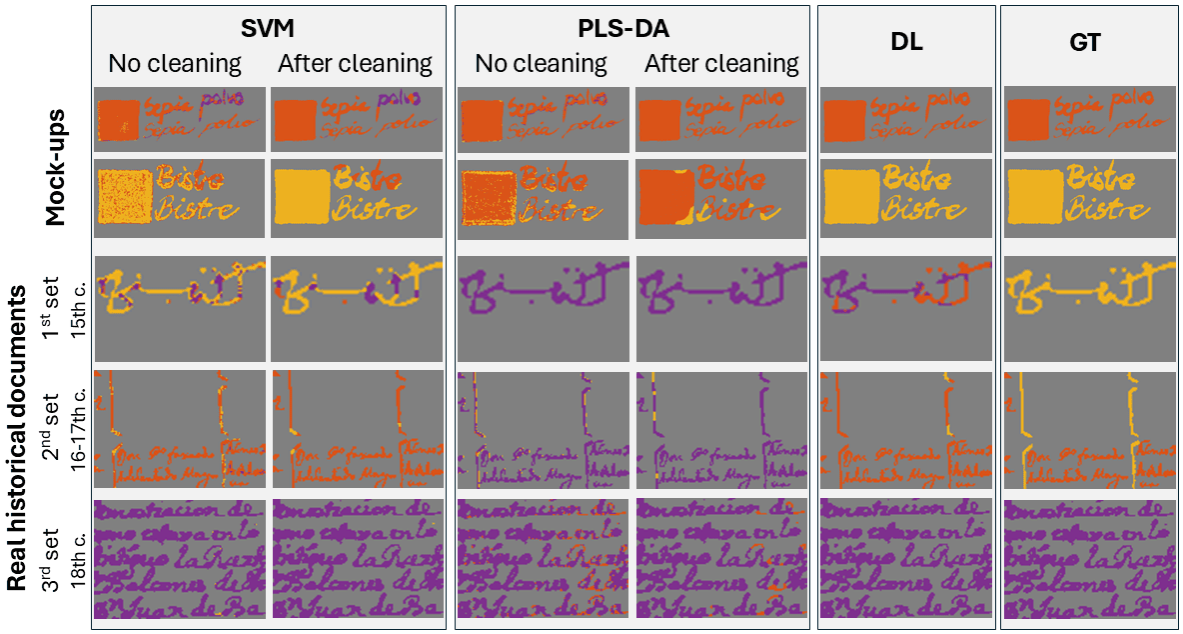

Figure 5.7: Examples of classification maps using the SVM model (columns 1 and 2), the PLS-DA model (columns 3 and 4), and DL (column 5). The Ground Truth (GT) images are shown in column 6. Purple: metallo-gallate ink (MGP); yellow: carbon-containing ink (CC); orange: non-carbon-containing ink (NCC).

struggles significantly, incorrectly classifying most pixels as NCC. The DL results (column five) are totally correct for the mock-up samples.

For the historical samples, the example in the 3rd row of the figure was difficult for most models. This sample, composed of CC ink (yellow color coding) on linen paper, is classified by SVM as containing all three ink classes, while PLS-DA incorrectly identifies it as MGP. DL mistakenly identifies the sample as a partial mixture of MGP and NCC: given the better performance observed qualitatively and quantitatively in other samples a possible explanation is that of the model having learned an incorrect bias by relying on stroke structure. Even after the cleaning post-processing is applied for the SVM and PLS-DA models, there are still some or all pixels that are misclassified. However, if the number of pixels classified into the three classes is considered, SVM provides a more accurate classification by correctly identifying

the majority of pixels belonging to the CC class. This classification challenge may be attributed to the sample's age, as the 15th-century manuscript exhibits ink fading due to aging, which increases the influence of the support on the final ink spectra and raises the reflectance (a wider explanation is given in subsection 5.3.4). In Figure 5.8, additional classification maps after post-processing are presented for all models studied. These maps illustrate the challenges in distinguishing two classes in the third column (16th–17th century sample), possibly due to the thin strokes, and the classification difficulties in the first-column CC sample, which may be related to its composition of 25% bone black and 75% sepia.

In the second historical sample (4th row), two types of ink can be found: a mixture of MGP and sepia (NCC) in the text, and CC ink in the braces. SVM and DL models have problems with the identification of carbon in the braces, correctly classifying only a few pixels, although correctly performing on the text. For the case of DL in particular, neural architectures for semantic segmentation are known to struggle on isolated thin structures, as typically demonstrated on pole lights in automotive applications [289]: this is due to a combination of learned neighborhood bias (which otherwise helps in correctly identifying large chunks of text) and neural structure limitations (already significantly improved by the DeepLabV3 architecture adopted in this work). PLS-DA misclassifies the entire sample as MGP ink, with only a few carbon pixels correctly identified in the braces.

Finally, the third historical sample (5th row), which is made entirely of iron gall ink, was correctly classified by the SVM and DL models, and nearly correctly classified by PLS-DA.

Figure 5.8: Classification maps obtained using all the models studied (SVM, KNN, LDA, RF, PLS-DA, and DL) after cleaning post-processing. The Ground Truth (GT) images are shown in the last row. Purple: metallo-gallate ink (MGP); yellow: carbon-containing ink (CC); orange: non-carbon-containing ink (NCC).

5.3.4 Case study: binarization and classification of inks in two full historical documents

In this section, two historical documents with higher complexity and size than the minicubes were tested using the best-performing traditional algorithm (SVM, according to subsection 5.3.3) and the DL model.

The hyperspectral data cube dimensions were [344×197×261] for the family tree document and [426×311×261] for the Arabic manuscript. The prediction times for both documents were 3.1 and 3.7 seconds, respectively, for the SVM model, demonstrating that it can provide near real-time classification once the spectral capture is performed, which is highly valuable for restorers, conservators, and researchers interested in the material composition of historical documents.

The DL model took 12.3 and 6.3 seconds for inference with a sliding window of 35×35 pixels, which increased to 30 and 15 seconds when accounting for data loading and transfer into GPU memory. This computational overhead is significant and should be considered for the implementation of any final application on systems with lower computational capabilities (see Table 5.8 for details on the computational environment).

In the family tree document (see Figure 5.9), the binarization step achieved good visual separation between the ink and support, with only some artifacts present in the lower right portion of the document along the right brace. A false RGB image was generated using the VNIR spectral bands at [605, 535, 430] nm. The SVM model successfully classified most parts of the text as NCC (orange color), while the DL model classified most of the pixels as MGP (purple color), and the remaining pixels as NCC. In addition, all carbon-based text was accurately located, although the DL model has some misclassified pixels as either NCC or MGP in the lower left and

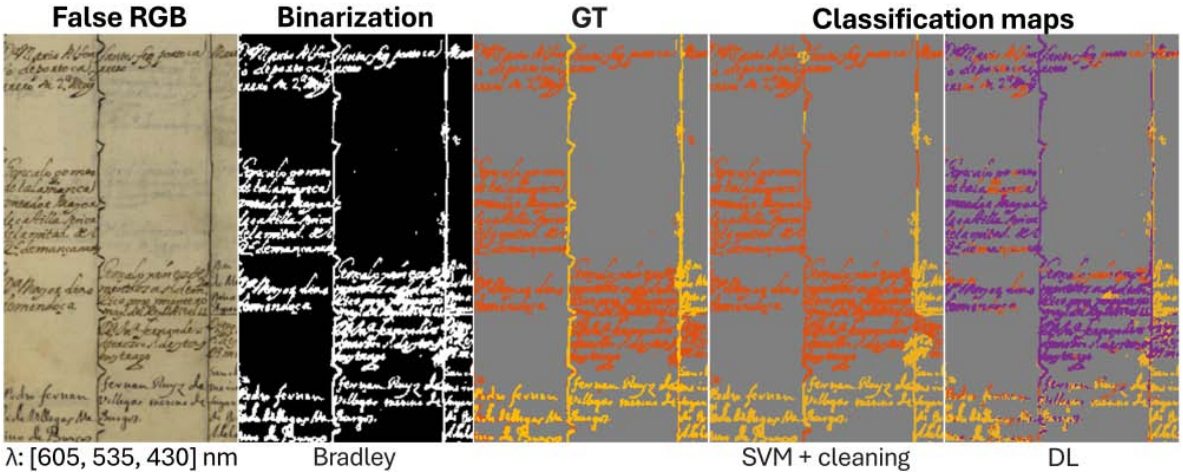

Figure 5.9: Family tree document. From left to right: false RGB image, binarization, GT, and classification maps using SVM model after cleaning post-processing and DL model. Predicted MPG pixels are shown in purple, NCC pixels in orange and CC pixels in yellow.

central parts. Some additional challenges arise due to the thin traces of the braces for both models, with SVM correctly identifying more pixels as CC ink (yellow color) in this part of the document. Overall, this document is classified more effectively by the traditional SVM-based model.

The analysis of the Arabic notarial manuscript (see Figure 5.10) presented more challenges. The binarization results were visually acceptable on the whole, with some bleed-through and stains in the upper and lower part of the document. The manuscript contains two types of ink: the main text and marginal note, both written with metallo-gallate ink with added earth (the correct class would be MGP), and the judge's validations and signature, which are composed of pure carbon ink (CC).

Compared to the family tree document, the classification results were less consistent, as both types of ink (MGP and CC for SVM or MGP and NCC for DL) were found in the same lines of text or words, which does not make sense in a document. However, after applying post-processing techniques for the SVM model, the visual results improved significantly. The judge's signature, located below the main text, is

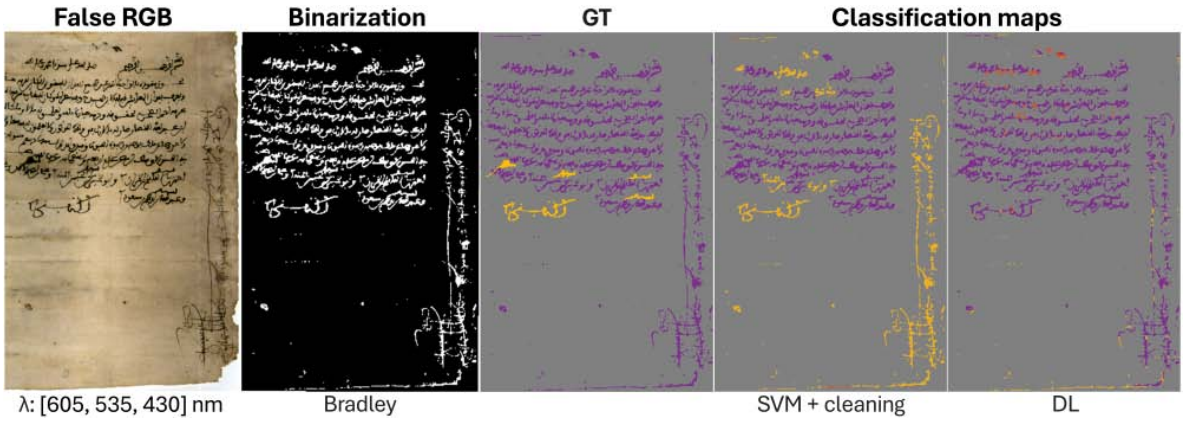

Figure 5.10: Arabic notarial manuscript. From left to right: false RGB image, binarization, GT, and classification maps using SVM model after cleaning post-processing and DL model. Predicted MPG pixels are shown in purple, NCC pixels in orange and CC pixels in yellow.

correctly classified as carbon-containing ink (CC, yellow color). This signature was included as well in the test set (see Figure 5.7), and it is misclassified as MGP (purple color) by the DL model. The main text is in most pixels correctly identified as MGP by both models. However, classification errors arise in other areas: the judge's validations are incorrectly labeled as MGP for both models, and the marginal note is mistakenly identified as CC for SVM, while correctly classified in most pixels by the DL model.

Further analysis of the document's reflectance in different regions (see Figure 5.11), reveals potential explanations for the misclassifications. When comparing the spectra, two distinct groups emerge: one containing the main text (MGP) and the judge's validations (CC), and another containing the signature (CC) and the marginal note (MGP). This explains the misclassification of the judge's validations as MGP and the marginal note as CC for the SVM model. Additionally, all inks become transparent in the SWIR range, complicating the classification further. Most of the mock-up CC training samples exhibit low reflectance in the SWIR range, which is crucial for accurately classifying samples in this class. The behavior of the inks

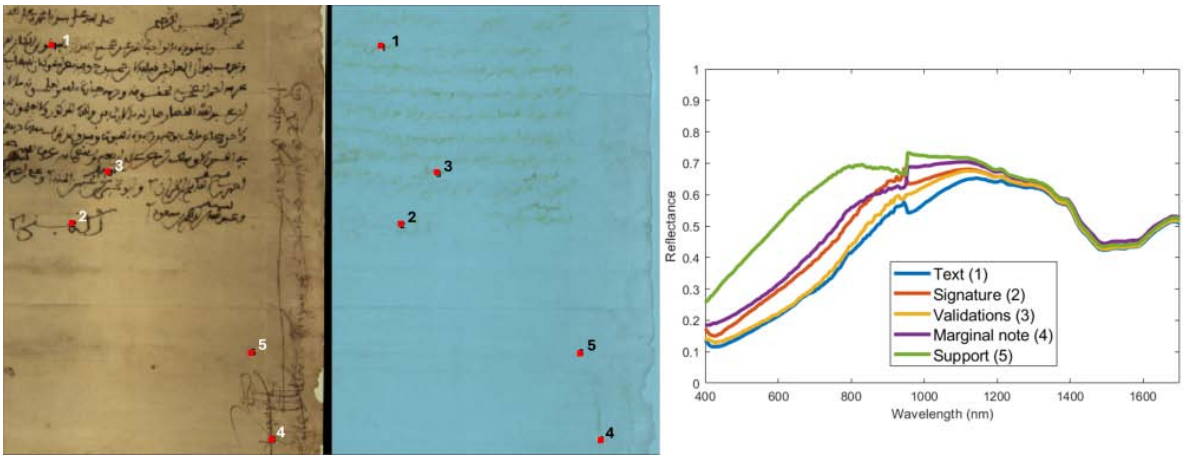

Figure 5.11: False RGB images of the Arabic notarial manuscript in the VNIR ([605, 535, 430] nm) and SWIR ([1300, 1100, 900] nm) spectral ranges. Regions of interest (1-5) were averaged to generate the spectral reflectance plot on the right.

in this document may be attributed to ink degradation, aging, and discoloration, which significantly alter the spectral properties of the ink and complicate classification. Similar spectral changes have been reported in previous studies, particularly in offset inks on paper subjected to artificial aging [290]. In contrast, the family tree document does not present these issues, likely due to better preservation and the fact that it is two centuries younger.

5.4 Conclusions

In this chapter, six classification models, including five traditional models (SVM, KNN, LDA, RF, and PLS-DA), and one DL-based model, were implemented for ink classification, and their performance was compared using both mock-ups and historical samples (test set), as well as two full pages extracted from historical documents (case study).

All studied models provided micro-averaged accuracy over 89.8% for the test set. The best results were obtained from the DL model, with micro- and macro-averaged

accuracy and recall above the 99% threshold. Nevertheless, among the traditional models, SVM emerged as the best option with all metrics above the 95% threshold and micro- and macro-averaged accuracy and recall above 97%. In both case studies, neither model achieved perfect results. The SVM misclassified fewer pixels and identified key features like the judge's signature in the Arabic notarial manuscript. This document was not included in the training of the model and presented notable challenges for accurate classification due to degradation, aging, and fading of CC inks in the SWIR range.

The choice between a traditional or a DL model can then be based mostly on the available computational resources and how pushing is the need for slightly better accuracy, since the training and hyperparameter tuning of the DL model require a considerable amount of processing resources and the prediction times for higher sized documents are longer. While traditional models could be trained and tested on a personal computer, the same machine could not tackle the training of the DL model. On the other hand, DL does not require a post-processing step that considers the spatial continuity of the classification maps, while traditional models benefit considerably from such post-processing.

The use of supervised classification models with HSI data has proven relevant for the material characterization of documents of historical interest. This can be related to the fact that reflectance imaging can provide indirect information about the molecular structure of the materials employed in the ink recipes, as highlighted in subsection 5.3.1 and mentioned in previous studies [291]. Unlike XRF mapping, reflectance imaging offers the capability to map both inorganic and organic materials or their mixtures. In this respect, it is important to consider data fusion of different spectral ranges as a pre-processing step to highlight distinctive features of the materials like fading in the SWIR range for MGP inks. A key limitation of the proposed

approach, compared to other analytical techniques, is the need for a large, annotated training dataset, which requires prior knowledge of the inks used in the documents. However, once the training phase is completed and the performance is evaluated with documents not included in the training set, this methodology eliminates the need for additional techniques to characterize new documents.

Although the identification of written areas in this study is achieved through binarization, this method may prove less effective in cases of poorly preserved texts or high variability, such as interference from complex backgrounds, fading and degradation of ink, stains on the paper, bleeding, paper transparency, or the presence of multi-colored inks. Future research could explore the use of automatic text zone identification schemes (e.g., bounding box-type approaches) or the integration of advanced DL architectures designed to handle these complexities and effectively separate text from the support.

Classification of inks in the Arabic notarial manuscript has been challenging due to spectral changes, which are likely associated with aging and discoloration. To address this issue, several strategies can be implemented: expanding training datasets with additional historical samples, though this is not always possible due to their fragility and restricted access imposed by conservation policies, and the use of unknown recipes in the materials present; using virtual aging simulations to model spectral shifts resulting from ink degradation; applying accelerated artificial aging to mock-ups in controlled environments (heat, humidity, and radiation) to study spectral changes, although this method may not fully replicate natural aging processes; and using microfading, which, while faster, is less comprehensive than artificial aging, as it only studies the effects of light exposure. These approaches could improve ink classification accuracy in historical materials.

The three classes used for this study provide very useful information for restor-

ers and historians interested in ink characterization of historical documents, since, for instance, MGP tends to show corrosion at the border of the trace, while CC will be more prone to fading. However, chapter 7 will focus on a more refined classification, aiming to distinguish the subclasses within the CC and NCC groups. Given that some inks are composed of multiple components, this will be addressed using spectral unmixing techniques. These methods can provide a more interpretable analysis of individual components and their concentrations in mixtures compared to DL or ML approaches. However, their effectiveness depends on the choice of mixing model, the accuracy of the extracted EMs (spectra of pure components), and the availability of a comprehensive reference library.

CHAPTER 6

Endmember extraction, unmixing and pigment identification in a copper plate painting

In this chapter, different EM extraction methods are compared in two stages of the same painting: a *Maternity* painted on a copper plate, before and after restoration. For the restored painting, two spectral hyperspaces are considered: spectral reflectance (R hyperspace) and the -log(R) hyperspace, in both the VNIR and SWIR regions. This analysis includes pigment identification and the detection of restored areas. The work presented in this chapter corresponds to specific objective O5 of this PhD thesis.

It extends the study published as:

E. M. Valero, M. A. Martínez-Domingo, **A. B. López-Baldomero***, A. López-Montes, D. Abad-Muñoz, and J. L. Vílchez-Quero, "Unmixing and pigment identification using visible and short-wavelength infrared: Reflectance vs logarithm reflectance hyperspaces," *Journal of Cultural Heritage*, vol. 64, pp. 290–300, 2023.

And the conference paper published as:

A. B. López-Baldomero*, M. A. Martínez-Domingo, J. Hernández-Andrés, R. Blanc, J. Vilchez-Quero, A. López-Montes, and E. M. Valero, "Endmember extraction for pigment identification pre- and post-intervention: A case study from a XVIth century copper plate painting," in *Archiving Conference*, vol. 20, pp. 198–203, Society for Imaging Science and Technology, 2023.

6.1 Introduction

Pigment identification in artworks remains a challenge despite numerous methodological advancements, particularly when relying exclusively on informa-

tion from non-invasive techniques. A universally accepted systematic approach has yet to be established for real pieces of artwork, as various factors complicate the process. Two of these factors are the condition of the piece and the presence of pigment mixtures.

1. The condition of the piece. Pigment identification is particularly relevant before a planned restoration or conservation intervention, to identify the materials employed by the artist or else previous conservation activities. But the piece often is covered by a patina of unclean varnish which strongly darkens and alters the reflectance of the pigments, especially for medieval or renaissance paintings, making accurate identification more difficult.

2. The presence of pigment mixtures. Pure pigments are seldom present in real artworks. In many situations, it is not practical to continue extending the number of reference pigments in the auxiliary patch collection until one covers all the possible mixtures that a given artist (often unknown) might have used. Then, spectral unmixing and EM extraction techniques can be potentially helpful. As discussed in subsubsection 2.5.1.3, these methods have been explored in previous studies with varying degrees of success [26, 156, 177].

This study presents a case in which both challenges are present and examines the effectiveness of spectral unmixing techniques for pigment identification. Grillini *et al.* [177] explored various mixing models, finding that the subtractive model outperformed others. However, the simplest way to transform a subtractive into an additive model (taking the -log of the spectral reflectance data) had not been explored previously. This transformation can be especially interesting if the EM extraction models used are linear, like NFINDR [169] or PPI [163].

Previous studies on pigment mapping and identification have predominantly relied on the VNIR range [134, 138, 161, 164, 168, 177, 292]. While some research has extended pigment mapping to the SWIR range [81,83,190], unmixing in this spectral region remains unexplored.

This study is divided into two main parts:

The first part serves as a preliminary study, where three EM extraction methods (NFINDR, NMF, and manual extraction) are compared in two stages of the same painting: before and after restoration. A subtractive mixing model is used in this phase, and pigment identification as well as the detection of restored areas are performed, focusing exclusively on the VNIR spectral range.

The second part evaluates the impact of transforming a subtractive mixing model into a linear one within the -log(R) hyperspace. The hypothesis is that linear EM extraction methods will perform better in this transformed space. To test this, three EM extraction methods, NFINDR, a DL-based algorithm and manual extraction, were applied in the reflectance (R) hyperspace, and two of them, NFINDR and manual extraction, were further tested in the -log(R) hyperspace, excluding the DL-based method in this space due to its non-linear nature. The study compares unmixing results across two spectral ranges (VNIR and SWIR) and two hyperspaces (R and -log(R)), using only the painting after restoration and a reference copper plate that contains some samples of pigments with the same support and preparation as the painting. The findings provide insights into the advantages and limitations of each method, spectral range, and hyperspace for pigment identification in historical artworks.

6.2 Preliminary study: Endmember extraction for pigment identification pre- and post-intervention

6.2.1 Material and methods

6.2.1.1 Painting on copper

The painting is on a somewhat unusual support (a copper plate) and shows a typical Maternity scene with the Virgin, an infant Jesus on her lap, and St. Joseph in the upper right side. The dimensions of the painting are 13.5 x 17.5 cm. An inscription on the back reads "Boceto di Pablo Veronese", suggesting a possible attribution. The artwork originates from a private collection in Spain and is currently being studied by a multidisciplinary team of experts in heritage conservation, art history, mineralogy, optics, computer science and analytical chemistry. The main goal of the investigation is to formulate a hypothesis regarding its authorship and establish its dating. This piece has a further interest because it has recently been submitted to a restoration intervention that has removed the previous darkened patina of badly applied varnish, covered some missing areas (lacunas), and applied a new varnish layer (Lefranc-Bougeois satined with UV protection) (see Figure 6.1 right). In the restoration process, the chromatic reintegration was made with Maimeri pigments ochre, natural earth, toasted ochre, Naples yellow (NY), Zinc white, Lapislazuli (LL), and Cadmium Red from Winsor & Newton. The visual appearance of the piece has noticeably changed after the intervention (see Figure 6.1).

The piece presents a simple palette with five main pigments: Bones Black (BB), Lead White (LW), Cinnabar (CN), NY and LL. Based on FRX and DRX analysis performed by the Department of Analytical Chemistry of the University of Granada

Figure 6.1: False RGB images (R = 605 nm, G = 535 nm and B = 430 nm) of the Maternity painting object of this study. Left: before intervention. Right: after intervention.

[293], CN can be detected in the Virgin's dress as well as the carnations, LL is present in both the background and the Virgin's mantle, LW is present in the carnations and the sleeves of the Virgin's chemise, NY can be detected in certain parts of the carnations and the Child's cloth, and BB is found in the background and shadowed areas. The relatively short number of pigments used makes this piece a good case study for demonstrating the possibilities of unmixing techniques to tackle pigment identification pre-intervention. Also, the opportunity to study it after restoration will highlight the possibilities of these techniques and determine if spectral imaging in general is a good tool to identify areas that have been intervened.

6.2.1.2 Reference copper plate

We aim to produce some evidence about the best approaches for automatic EM extraction and pigment identification from the list of EMs using an auxiliary copper plate with pigments applied on the same support.

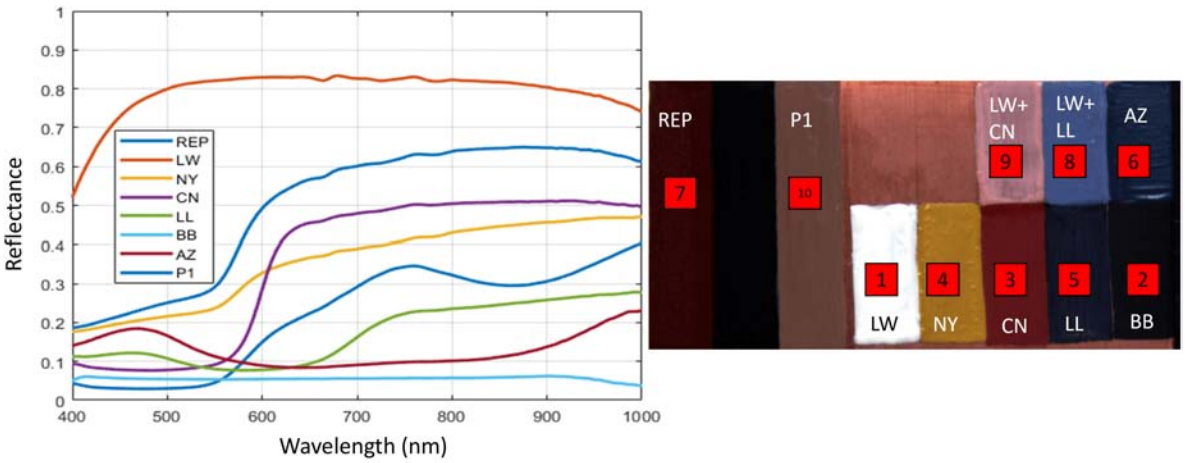

Figure 6.2: (Left) Spectral reflectance of manually extracted 30x30 pixel areas containing the different pigments and mixture deposited on the auxiliary copper plate (right).

The preparation used for the reference copper plate is in accordance with documented period techniques from the XVIth century [294–297]. The unpolished copper plate (1 mm thickness) was cleaned with calcium carbonate and vinegar and rubbed with ground garlic. Then, a layer of $CaSO_4$ powder bound with linseed oil was applied, followed by a preparation layer (P1) of BB mixed with Read Earth (REP) and LW, also bound with linseed oil. Then, the pigments and mixtures are deposited on the prepared surface. There are seven pure pigments: LW, NY, BB, CN, LL, Azurite (AZ) and REP. Three more patches with mixtures are added: LL + LW, CN + LW, and the P1 mixture. The materials applied include preparation layers and pigments found in the painting [293], as well as some that are not present (such as AZ). The pigments (Kremer Pigmente GmBH) were bound with linseed oil and applied with a brush.

After capturing the reference plate (described in subsubsection 6.2.1.3), a reference library of eight spectra for pigment identification was built using the average reflectance of 30x30 pixels areas within the patches labeled as 1-7 and 10 in Figure 6.2 right. The spectral reflectances included in the reference library are shown in

Figure 6.2 left.

6.2.1.3 Hyperspectral image capture

The VNIR spectral images of the reference copper plate and the two versions of the painting were captured and processed as described in subsection 3.2.2 and subsubsection 3.2.3.1. The SWIR range was excluded in these preliminary tests, as VNIR had been used in previous unmixing studies for pigment detection and allowed the use of color information for evaluation.

6.2.1.4 Endmember extraction and pigment identification

Once the spectral images were captured, three methods for EM extraction were used: NFINDR [169], NMF [170] and manual extraction of spectra in selected areas from the painting that appeared to contain each of the five pigments in its most pure form, although this was not possible for the NY pigment, which appeared exclusively in mixed form.

Three sets of EMs are extracted from the spectral images of the painting, and unmixing is carried out using an optimization process with a cost function based on a combined metric formed by two components: the complement of the Goodness-of-Fit coefficient ($cGFC$, [298]) and the $RMSE$ [177]. $cGFC$ is sensitive to shape changes, while $RMSE$ is sensitive to both changes in scale and shape when comparing two spectra. A perfect match would have zero $RMSE$ and $cGFC$ values. The final form of the metric is:

$$M = cGFC + \beta \cdot RMSE \tag{6.2.1}$$

β is a scaling parameter to balance the contribution of both sub-metrics to the fi-

nal combined metric. In a preliminary experiment, the optimal value for β parameter to ensure equal contribution of both metrics was found to be $\beta = 1.0936$.

The fmincon function with the interior point algorithm [299] implemented in MATLAB® was used for optimization, with the sum-to-one constraint ($\sum_{i=1}^{q} \alpha_i = 1$) and a lower bound for the weights α_i of zero. The mixing model is a subtractive model [300], where the spectral reflectance of *i* EMs are multiplied consecutively, elevated to the power of the concentrations (Equation 6.2.2) for each wavelength.

$$\Upsilon = \prod_{i=1}^{q} \rho_i^{\alpha_i} \tag{6.2.2}$$

Where Υ is the spectral reflectance of the mixture, *q* is the number of candidate EMs, ρ_i is the spectral reflectance of the *i*th EM, and α_i its concentration.

Both NMF and NFINDR have the limitations of a linear mixing model assumption, while subtractive mixing has been proven to be performing best for pigments on canvas [177]. On the other hand, the manual EM extraction from the painting has the limitation of not corresponding to pure pigments and being affected by aging. The unmixing process will estimate the weights in the mixture, and from them, we obtain the concentration maps and the error maps of the five EMs (see Figure 6.4). This will allow us to determine which EM set is the best for the spectral reconstruction of the painting using EMs and estimated weights. This will be evaluated using both spectral and colorimetric differences on a pixel-by-pixel basis ($cGFC$, $RMSE$ and $CIEDE00$) [298, 301].

Finally, we will use the auxiliary set of reference spectra for trying to find the pigment that presents the closest spectral distance to each EM for the three EM extraction procedures. This spectral distance is calculated using a combined distance metric between the pigments in the plate and the EMs obtained:

$$MId = cGFC + 0.5MSE + 0.02\Delta E_{00} \tag{6.2.3}$$

The coefficients are obtained using the following tolerances for the three components of the metric: 0.01 for $cGFC$, 0.02 for MSE and 0.5 for ΔE_{00}. Those coefficients ensure all the factors would contribute equally to the metric value in an acceptable reflectance match. The tolerance values are based on experience and correspond with tolerances found in the literature for spectral estimation [298]. The metric contains a color difference term because color can be relevant for pigment identification, while both scale and shape differences in spectra are also accounted for. After computing this metric, the label corresponding to the reference pigment with the minimum MId metric value is assigned to each EM. Finally, the hit rate is calculated as the percentage of correctly identified pigments compared to the total number of pigments present in the painting.

The process will be repeated for the restored painting, allowing us to see if there are any differences in the performance of the unmixing algorithm, and if the renovated areas can be identified from the concentration maps. All the calculations have been performed using MATLAB® (Release 2022a, The MathWorks, Inc., Natick, MA, USA). The NFINDR and NMF algorithms used are those provided by MATLAB®, and the *fmincon* function is used for the unmixing optimization.

6.2.2 Results

6.2.2.1 Concentration maps and endmember sets

The EMs sets differ for the three EM extraction methods, and are not directly similar to the reference pigments (obtained with the auxiliary copper plate) in any of the cases, similarly to the findings in [302]. In Figure 6.3, the EMs extracted from the

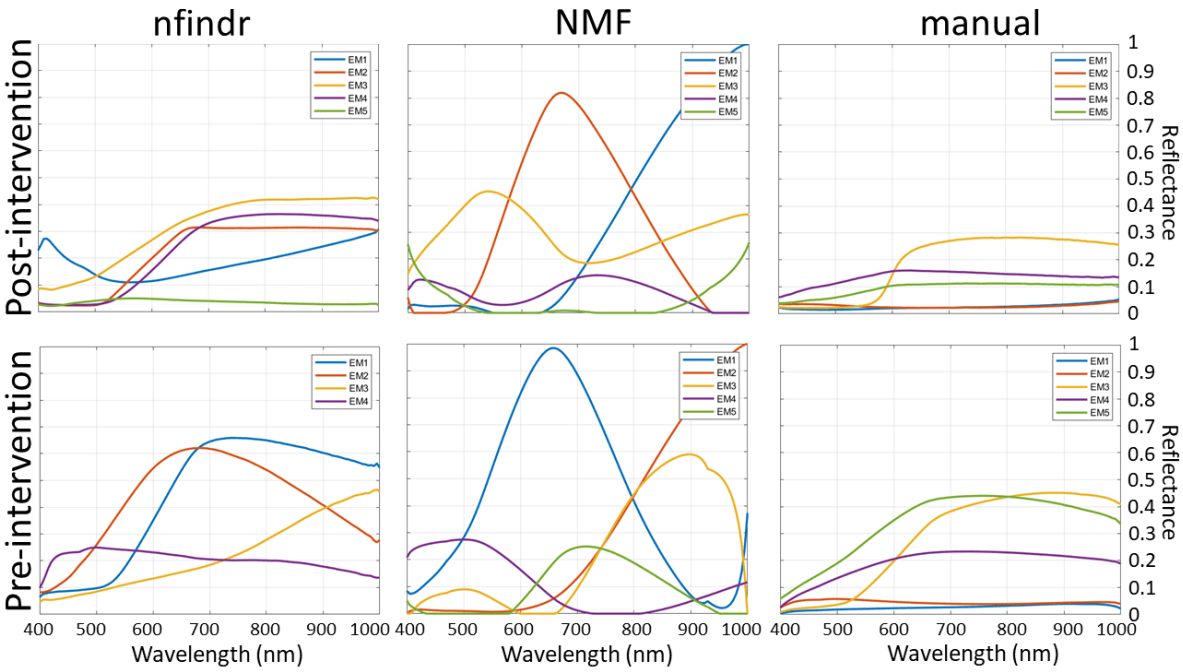

Figure 6.3: EM libraries extracted using NFINDR (left), NMF (center) and manual extraction (right) methods pre- (lower row) and post-intervention (upper row).

painting pre- and post-intervention are shown. The NMF algorithm implemented in MATLAB® provided EM reflectances with maximum values over 1, so they were normalized by the maximum value. The spectral reflectances are flatter for the manually extracted library than for those extracted with NFINDR or NMF algorithms. In general, post-intervention EM reflectances show lower values than pre-intervention reflectances. This could be due to the removal of the whitish patina present in the original painting during the restoration process. NMF seems to provide reflectance curves that differ more from real pigments, so the preliminary hypothesis is that NFINDR and manual library will work best for the concentration vector and spectral estimation part of the unmixing process.

Pre-intervention

In Figure 6.4, concentration maps for the EM most similar to CN in the three sets

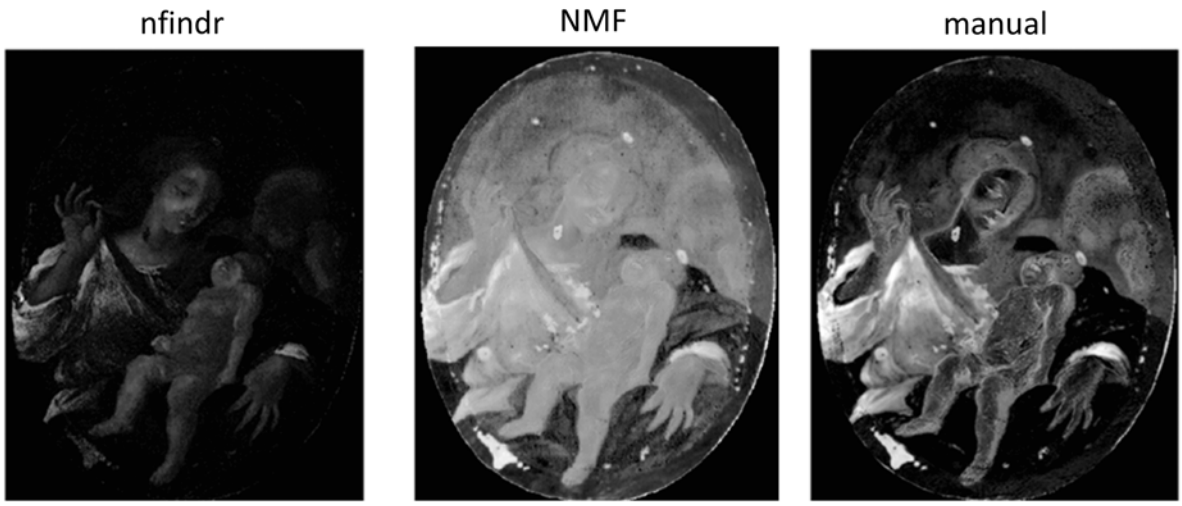

Figure 6.4: Concentration maps for the EM most similar to CN in the pre-intervention painting. Left: NFINDR. Center: NMF. Right: manual extraction.

are shown for the unrestored painting (EM1 in NFINDR, EM2 in NMF, and EM3 in the manual extraction set). It can be seen that manual extraction method identifies the copper support as CN. NMF detects CN in the background, not only the dress and the carnations.

According to the concentration maps (not shown here for brevity), EM3 obtained with NMF seems to represent the LL pigment, as well as EM4 in NFINDR (although this is less clear) and EM2 in the manual extraction. EM4 and EM5 in the manual extraction could represent the LW pigment, as well as EM2 in NFINDR, and EM1 in NMF (although it appears all over the painting).

Post-intervention

In Figure 6.5, concentration maps for the EM most similar to CN in the three sets are shown for the restored painting (EM4 in NFINDR, EM2 in NMF, and EM3 in the manual extraction set). Similar to the results obtained in the pre-intervention painting, the NMF method detects CN in the background. The manual and NFINDR concentration maps seem plausible to the CN pigment, finding a higher concentra-

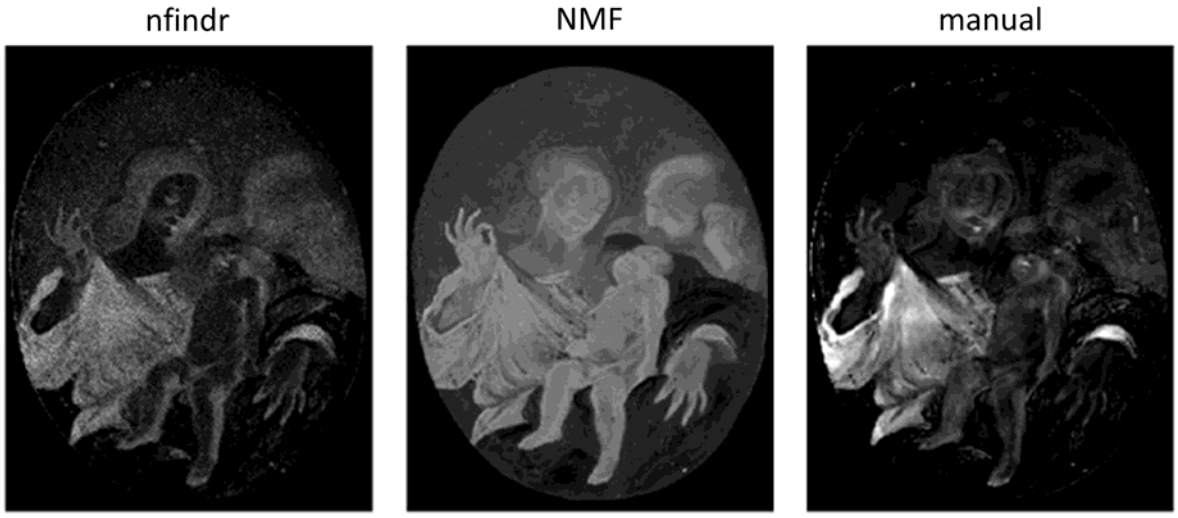

Figure 6.5: Concentration maps for the EM most similar to CN in the post-intervention painting. Left: NFINDR. Center: NMF. Right: manual extraction.

tion in the dress with the manual extraction method. In this state of the painting and according to the concentration maps, EM4 obtained with NMF seems to represent the LL pigment, as well as EM1 in NFINDR and EM2 in the manual method.

6.2.2.2 Spectral estimation quality

The $RMSE$ error maps calculated from the differences between the reconstructed spectra using Equation 6.2.2 and the spectra from the spectral image of the unrestored painting are shown in Figure 6.6. Note that the range of values for $RMSE$ is restricted to [0,1] for reflectance data. The mean $RMSE$ value shows that NFINDR fails to provide an accurate spectral reconstruction, while the manual extraction produces the best results.

The $cGFC$ error maps for the restored painting are shown in Figure 6.7. We can see that NMF method fails to reconstruct the shape of the spectra compared to manual extraction and NFINDR.

In Table 6.2, the $cGFC$, $RMSE$ and ΔE_{00} values obtained from the comparison

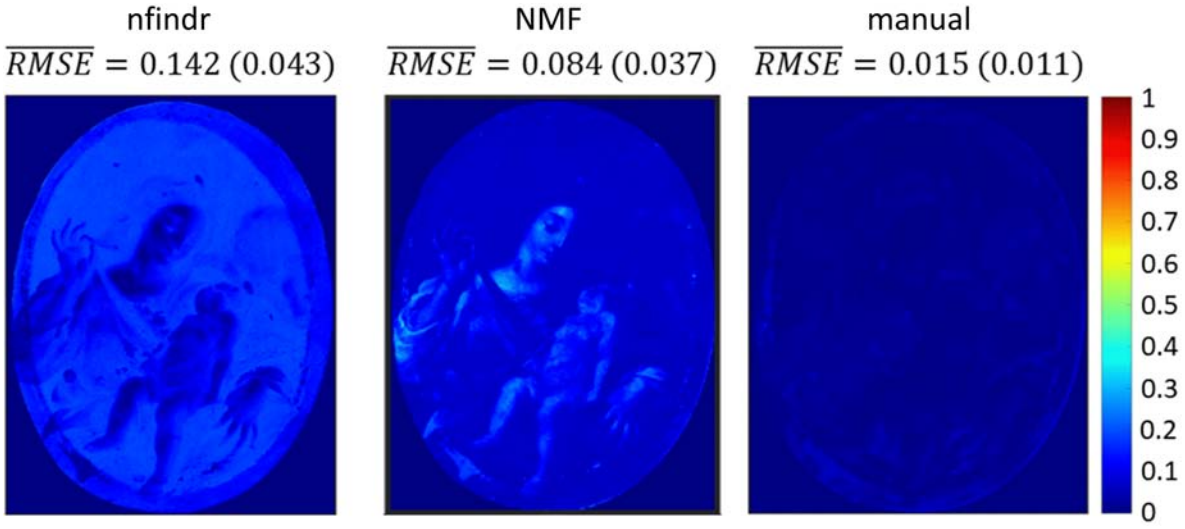

Figure 6.6: $RMSE$ error maps for the pre-intervention painting. Mean (and standard deviation) values are shown on top of each map. Left: NFINDR. Center: NMF. Right: manual extraction.

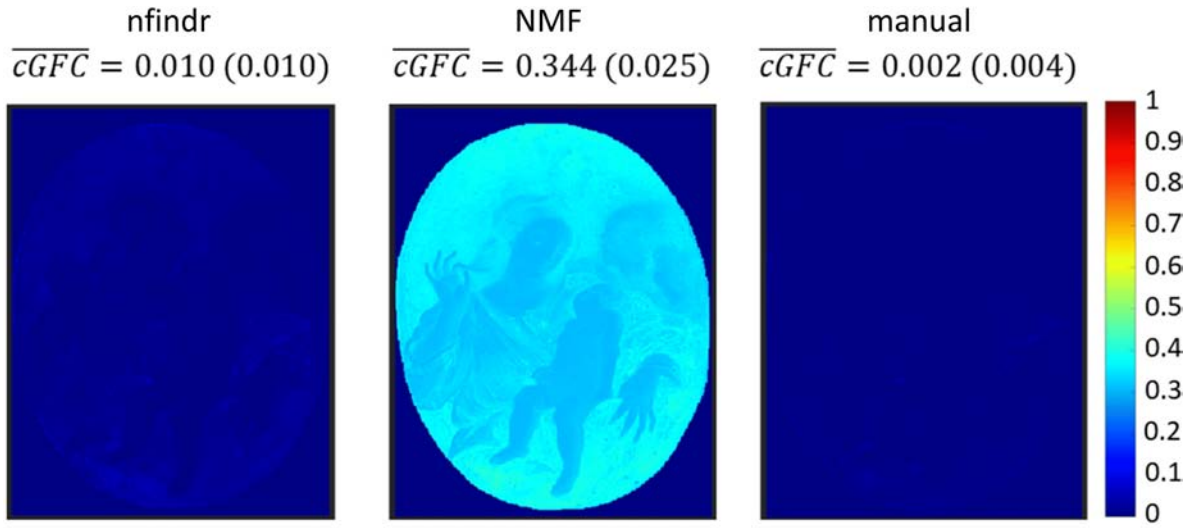

Figure 6.7: $cGFC$ error maps for the post-intervention painting. Mean (and standard deviation) values are shown on top of each map. Left: NFINDR. Center: NMF. Right: manual extraction.

between the estimated and the original spectra are shown for each EM set and state of the painting. The best $cGFC$, $RMSE$ and ΔE_{00} values are obtained in both states for the manual extracted library. Comparing the painting pre- and post-intervention, lower values for $cGFC$ are found for the pre-intervention painting with the NMF

Table 6.1: Spectral estimation quality metrics for the three EM sets pre- and post-intervention.

	EM set	cGFC (STD)	RMSE (STD)	DE00 (STD)
Pre-intervention	NFINDR	0.007 (0.005)	0.142 (0.043)	19.72 (6.28)
	NMF	0.296 (0.030)	0.084 (0.037)	23.87 (5.34)
	Manual	0.005 (0.008)	0.015 (0.011)	3.94 (2.52)
Post-intervention	NFINDR	0.010 (0.010)	0.074 (0.031)	12.64 (5.12)
	NMF	0.344 (0.025)	0.073 (0.023)	29.38 (3.28)
	Manual	0.002 (0.004)	0.014 (0.0015)	4.17 (2.29)

and NFINDR libraries. Comparing the $RMSE$ values, post-intervention painting provides the lowest values for all three quality metrics. For the ΔE_{00}, the values depend on the EM extraction method used, not so much influenced by the painting state. For the post-intervention painting, the mean $cGFC$ value and the error map (Figure 6.5 lower row) show that NMF is the worst algorithm in providing accurate spectral reconstruction. Comparing the results with manual and NFINDR extraction methods, the manual method is 5 times better than NFINDR in terms of $cGFC$. Our preliminary hypothesis, i.e., that NFINDR and manual libraries will work better than the NMF library, is true for the $cGFC$ and the ΔE_{00}, but not for the $RMSE$. This means that the NMF algorithm introduce some additional changes in shape in the estimated spectra with respect to the original, but overall, it is better than NFINDR at capturing the scale of the spectral reflectances. This could be because NFINDR is more sensitive to the problem of using a linear mixture model for EM extraction than NMF.

6.2.2.3 Pigment identification

Pigment identification results are shown using the reflectance of the pigments in the auxiliary copper plate as reference (Table 6.2).

Depending on the state of the painting, different results have been obtained with

Table 6.2: Pigment identification results for the three EM sets pre- and post-intervention.

	EM set	Assigned labels	Hit rate	MId range
Pre-intervention	NFINDR	P1, REP, NY	20	0.29-0.53
	NMF	P1, REP, LL, AZ	20	0.61-0.79
	Manual	BB, REP, NY	40	0.14-0.40
Post-intervention	NFINDR	LL, REP, NY, REP, BB	60	0.13-0.27
	NMF	LL, P1, LW, BB, AZ	60	0.506-0.875
	Manual	BB	20	0.11-0.42

the three libraries. For the pre-intervention painting, the best result is obtained with the manual library with a hit rate of 40%, compared to the 20% obtained with NFINDR and NMF libraries. Even being the best, it fails to identify CN, LW and LL. Referring to the post-intervention painting, the best results are obtained with NFINDR and NMF libraries, with a hit rate of 60%. Both identify LL and BB, NY is identified in NFINDR, and LW is identified in NMF. AZ is identified in NMF, but it is not present in the painting. In both states of the painting, NMF library shows the higher MId values. Compared to the other libraries, the manual extracted library is inherently disadvantaged because the reflectances of the paintings are consistently much lower than those of the reference library. As a result, more than one EM is classified as BB, which does not happen with the other libraries. These results are conditioned by the restricted and specific set of pigments used in the copper plate, which contains seven pigments and three mixtures.

6.2.2.4 Identification of re-painted areas

First, re-painted areas were searched from the concentration maps of the restored painting. They seemed to appear in the EM1 of the NMF library (Figure 6.8 left) in the form of small white spots, and in the EM2 of the same library and the EM3 of the manual extracted library as black spots. The restored areas were not easily visible in

Figure 6.8: Concentration map for the EM 1 obtained with NMF (left). False RGB image of the restored painting with channels [695,980,905] nm (right).

the other concentration maps and libraries.

Due to the slightly different color showed by the re-painted areas, we decided to search the best three bands of the spectral cube that provided the higher F1-score value (see Equation 2.5.4) comparing the false RGB image to a GT of re-painted areas obtained manually from the original painting. The segmentation of re-painted areas was performed in the false RGB color space, looking for the optimum minimum and maximum values for the three channels. In a preliminary study, we performed a band-by-band intensity thresholding, but the results were worse than those obtained with a color-based segmentation. The optimization was done with the functions surrogateopt and genetic algorithm, implemented in MATLAB®.

For the surrogateopt function, the best combination of three bands was [410,965,705] nm. The minimum and maximum RGB values were: [(75,94), (17,162), (23,129)] for red, green, and blue respectively, and the best F1-score value was 0.112. For the GA, the best combination of three bands was [695,980,905] nm (Figure 6.8

right). Re-painted areas can be seen with a light green color against the background. The minimum and maximum thresholds in RGB were: [(36,91), (133,175), (106,194)], and the best F1-score value was 0.170. The best results were obtained with GA, although they were not good for any of the optimizations. This makes sense because GA is a global optimizer while surrogateopt attempts to find the global minimum using few objective function evaluations, which can lead to a local instead of a global solution. The main advantage of surrogateopt is the reduced computational cost and execution time.

To improve the results obtained in the detection of re-painted areas in a restored painting with VNIR information, it could be beneficial to use the SWIR range, since this type of radiation can penetrate more deeply into the different layers of the painting.

6.2.3 Discussion and conclusions

In this section, three EM extraction methods (NFINDR, NMF and manual extraction) are compared in two stages (pre- and post- intervention) of the same painting. The spectral images of the painting were captured from 400 to 1000 nm. Pigment identification was also conducted using an auxiliary set of reference spectra on the same support and with the same preparation, but corresponding to new materials.

The EMs obtained from the manual extraction method were flatter than those obtained with NFINDR and NMF methods. The reflectance curves of the NMF EMs differed from real pigments. The best concentration maps (i.e. those most similar to the real pigment distributions) were obtained with the manual extraction method.

Attending to the spectral reconstruction of the three EM extraction methods, manual extraction performed the best, with the lowest values of $cGFC$, $RMSE$ and ΔE_{00} for both states of the painting. Comparing the painting pre- and post-

intervention, lower values for $cGFC$ were found for the pre-intervention painting with the NMF and NFINDR libraries, but $RMSE$ was lower in the post-intervention case. This can maybe be linked to the fact that the renovated painting is more inhomogeneous than the original painting, and the spectral reconstruction is harder for this sample. The $RMSE$ values can be explained because the highest reflectance values correspond to NMF, and this is an initial advantage to get the scale right in the spectral reconstruction. NFINDR and manual libraries worked better than the NMF library considering the $cGFC$ and the ΔE_{00}, but not for the $RMSE$.

The results of pigment identification were influenced by the state of the piece. Manual extraction performed the best for the unrestored painting, identifying only two pigments, while NFINDR and NMF performed the best for the restored painting, identifying three over five pigments. The manually extracted EMs differ more from the reference library for the pre-intervention painting than for the restored one. This explains the overall worse identification results for the first case. However, restoring the piece enhances the efficiency of automatic algorithms in extracting EMs for pigment identification (not for reconstruction of the spectra, which is inherently more difficult). This highlights the disadvantage of the manual extracted library compared to the automatically extracted ones. Even with a very reduced palette of pigments, the results were not as expected. This could be due to differences between the spectra of the reference pigments and the actual pigments present in the painting. These findings underscore the importance of having an appropriate auxiliary palette of reference pigments for pigment identification.

The comparison of the information present in the concentration maps obtained from the unmixing process does not appear to be sufficient for detection of repainted areas. Even after finding the triplet of spectral bands that produced a more salient visualization of the restored areas with an optimization approach, the results

obtained by thresholding were not similar enough to the ground truth image containing all the restored areas. More sophisticated segmentation algorithms to detect re-painted areas can be considered in future work, as well as information in different spectral ranges.

The limitations imposed by EM extraction algorithms based on linear unmixing may be addressed by using a non-linear model to approximate the reflectance of mixed pigments. This approach is explored in section 6.3.

Spectral image capture is totally non-invasive, offers high spatial resolution, and significantly less time for capture and analysis than alternative techniques. These advantages highlight the importance of devoting more effort to refining unmixing methods so that they can perform well in EMs extraction and pigment identification.

6.3 Unmixing and pigment identification using visible and short-wavelength infrared: Reflectance vs logarithm reflectance hyperspaces

6.3.1 Material and methods

6.3.1.1 Samples

Two main objects were used in this study: an auxiliary (reference) copper plate (whose preparation is described in subsubsection 6.2.1.2), from which a checkerboard image was extracted, and the painting on copper with the inscription "Boceto di Pablo Veronese" on the back.

Ten 32x32 pixel areas were extracted from the spectral image of the reference copper plate (see Figure 6.2 right) to build the reference checkerboard image shown

Figure 6.9: (Left) Restored painting with the 3×3 pixels areas used for manual EM extraction marked in bright yellow. (Right) Reference checkerboard image.

in Figure 6.9 (right).

The painting used in this section was the same one described in subsubsection 6.2.1.1, after the restoration process. An RGB image obtained with three spectral bands (R = 605 nm, G = 535 nm and B = 430 nm) is shown in Figure 6.9 (left), with several 3x3 pixels areas marked in bright yellow, which were used to build the manual extraction (MEx) EM library as explained in subsubsection 6.3.1.3.

6.3.1.2 Spectral image capture and reference library

The VNIR and SWIR spectral images of the reference copper plate and the painting were captured and processed as described in subsection 3.2.2 and subsubsection 3.2.3.1.

After capturing the reference plate, a reference library (REFL) was built as described in subsubsection 6.2.1.2. All image processing and unmixing modelling has been performed using MATLAB® software (Release 2022a, The MathWorks, Inc.,

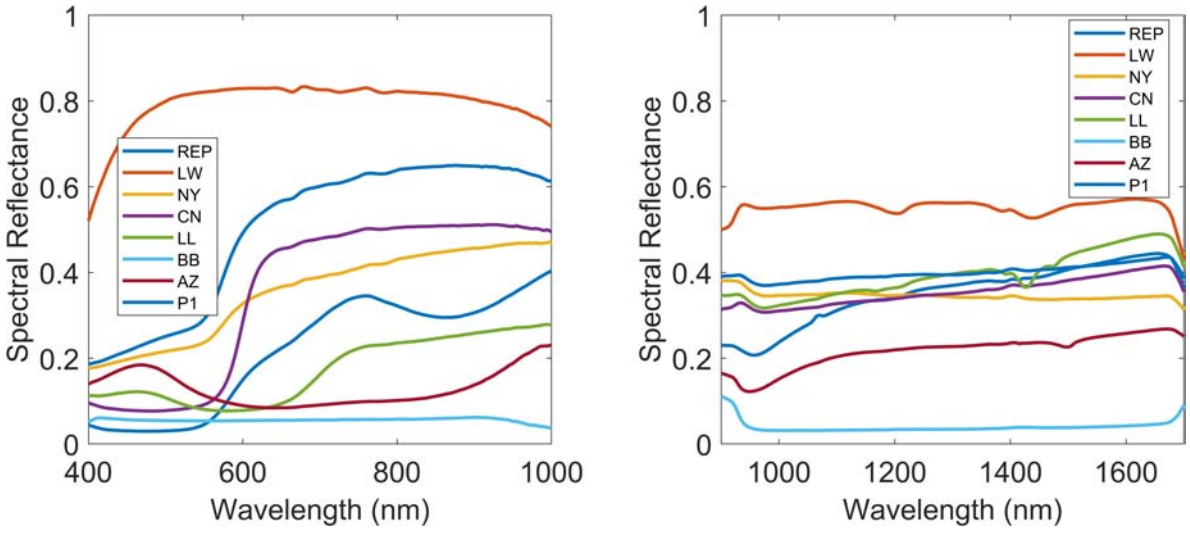

Figure 6.10: (Left) REFL spectra in the VNIR range. (Right) REFL spectra in the SWIR range.

Natick, MA, USA). In Figure 6.10, the spectral reflectances from the REFL in both VNIR and SWIR ranges are shown. The spectra differ both in shape and scale.

6.3.1.3 Unmixing methods

The process of unmixing often takes two steps: EM extraction and concentration estimation.

Concentration vector estimation and reflectance hyperspaces

Two mixing models have been used depending on the hyperspace. In the R hyperspace, the subtractive model [300] was used (see Equation 6.2.2).

Some blind EM estimation algorithms, like NFINDR [169] and FIPPI (Fast Iterative Pixel Purity Index) [303], are based on a linear model. This fact prompted the idea of introducing the -log(R) hyperspace for performing unmixing. In -log(R) hyperspace, the spectral reflectance of a subtractive mixture (see Equation 6.2.2) transforms into a linear combination of EMs for each wavelength, as shown in Equation 6.3.1 [304].

$$\Upsilon = \sum_{i=1}^{q} \alpha_i \rho_i \tag{6.3.1}$$

The goal of spectral unmixing is retrieving the vector of concentrations ($C = (\alpha_1, \alpha_2, \ldots, \alpha_q)$) from the spectral reflectance of the mixture (Υ) and the library of candidate EMs ($E = (\rho_1, \rho_2, \ldots, \rho_q)$). For this, a constrained optimization method is used, as described in subsubsection 6.2.1.4.

Summarizing, both R and -log(R) hyperspaces, each with its unmixing model (subtractive and additive, respectively), are used for the checkerboard image with different EM libraries in VNIR and SWIR ranges. The best performing method according to the criteria explained in subsubsection 6.3.1.5 is selected for analyzing the painting on copper.

Endmember extraction

EMs can be obtained by blind separation using different algorithms, like PPI [305,306], or FIPPI [303] and NFINDR [169]. The NFINDR and DeepGun [171] methods have been selected for this study as representative instances of automatic EM extraction methods with different design strategies. The NMF method was discarded after the preliminary study, as it produced EMs that differed from the reflectance spectra of real pigments. The MATLAB® libraries of ENVI integrated package for NFINDR, and the code provided by the authors of DeepGun [307] was used with the default parameters adapted to the number of extracted EM. For DeepGun, the extraction was performed only in R hyperspace, because the DeepGun algorithm is non-linear in this hyperspace.

Other possibility is to extract the library directly from the painting, out of representative areas including mixed pigments (Manual Extraction or MEx method) [138]. Five 3x3 pixel representative areas (shown in Figure 6.9) of black, blue, red, white,

and yellowish colors were extracted from the image of the painting. The spectra were averaged to build the MEx_p library. The MEx_p library is very likely constituted by mixtures and not pure EMs, but our hypothesis is that this library will provide more accurate concentration estimations since it is obtained directly from the painting. In any case, the MEx_p EMs will be used for pigment identification using the REFL spectra shown in Figure 6.10 as reference.

Summarizing, we have seven libraries in R hyperspace for each spectral range. Four of them are extracted from the copper reference board: REFL, which will only be used for pigment identification and other three (with 7 EM), which will be used for unmixing: the NFINDR library (NFDL), the DeepGun library (DeGu), and the Manual extraction library (MEx), which includes the same pure pigments as REFL but extracted from different areas. The remaining three libraries (with 5 EM) are extracted from the painting spectral image: NFDL_p, DeGu_p and MEx_p. The goal is not to find the best extraction method in absolute terms, but to choose three representative instances of extraction methods to showcase the proposed methodology. In -log(R) hyperspace, for each spectral range we have the two NFINDR libraries with 7 and 5 EMs, and the two M_Ex libraries for concentration estimation.

6.3.1.4 Pigment identification

The method consists of two parts: the first, computation of a combined distance metric between each pair of spectra from the candidate library and REFL. The metric for the VNIR spectra is presented in Equation 6.2.3. The combined metric in the SWIR range contains only the two first factors. as explained before, the lower the MId value, the higher the similarity between the compared spectra.

The second part is the label assignment: the REFL label corresponding to the pigment with the minimum metric value is assigned to each EM. Finally, the hit

rate of the pigment identification process is calculated as the percentage of correctly identified pigments.

6.3.1.5 Evaluation of results

The evaluation of the results obtained is based on three factors:

a) Spectral reconstruction: the similarity between estimated spectra, obtained by using the forward mixing model with the EMs and the estimated concentrations, and the original spectra, assessed separately by the three metrics that form MId (Equation 6.2.3).

b) Visual assessment: using concentration or presence maps. The data shown in Ref. [293] will be used to determine if the maps are plausible for the painting.

c) Hit rate: from pigment identification.

Figure 6.11 shows the workflow of the methods to clarify the procedures described in the previous subsections.

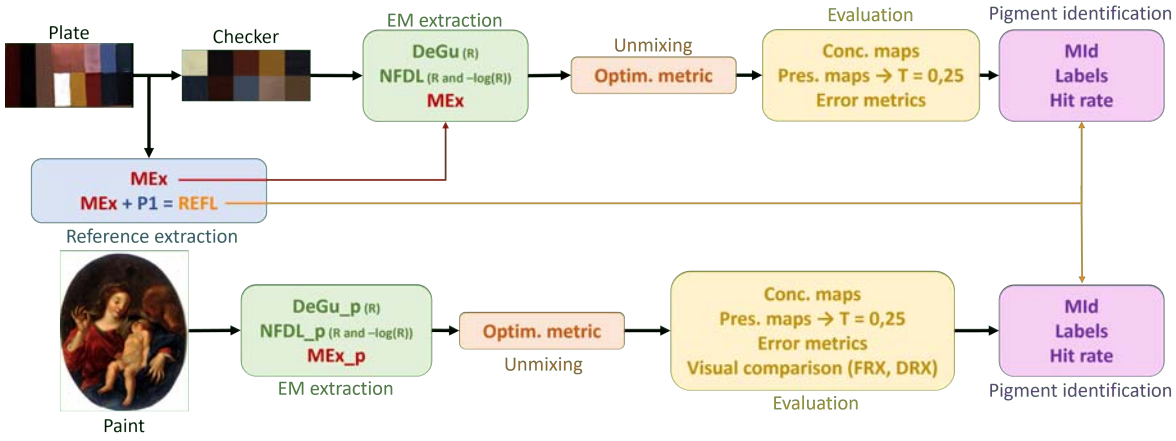

Figure 6.11: Workflow of the steps used in the different phases of this study.

6.3.2 Results and discussion

6.3.2.1 Copper reference plate

In this section, EM extraction and unmixing results for the spectral data of the checker reference image (Figure 6.9 right) are shown.

Unmixing results

Extracted endmember libraries

In Figure 6.12, the three EM libraries with seven EM each extracted from the checkerboard reference image in R and -log(R) hyperspaces are shown for the VNIR range. The libraries in -log(R) hyperspace were transformed back into R hyperspace after extraction, to facilitate comparison between hyperspaces. For manual extraction, there is only one EM library shown in Figure 6.12, since the reflectance values obtained would be the same after extraction in -log R hyperspace and applying the inverse transform to R hyperspace. There is a noticeable difference between some of the EM of the NFINDR libraries extracted in the two hyperspace (Figure 6.12 left column). The similarity between the MEx Library and NFINDR is higher than for DeepGun, which has lower maximum reflectance values for the EM with highest reflectance.

In Figure 6.13, the corresponding libraries in the SWIR range are shown. The observation regarding the similarity between the MEx and NFINDR libraries is also valid for the SWIR range. The DeepGun EMs tend to present less variety of shape and also lower maximum reflectance values in this range.

Spectral reconstruction quality

As explained in subsubsection 6.3.1.5, the similarity between the captured spec-

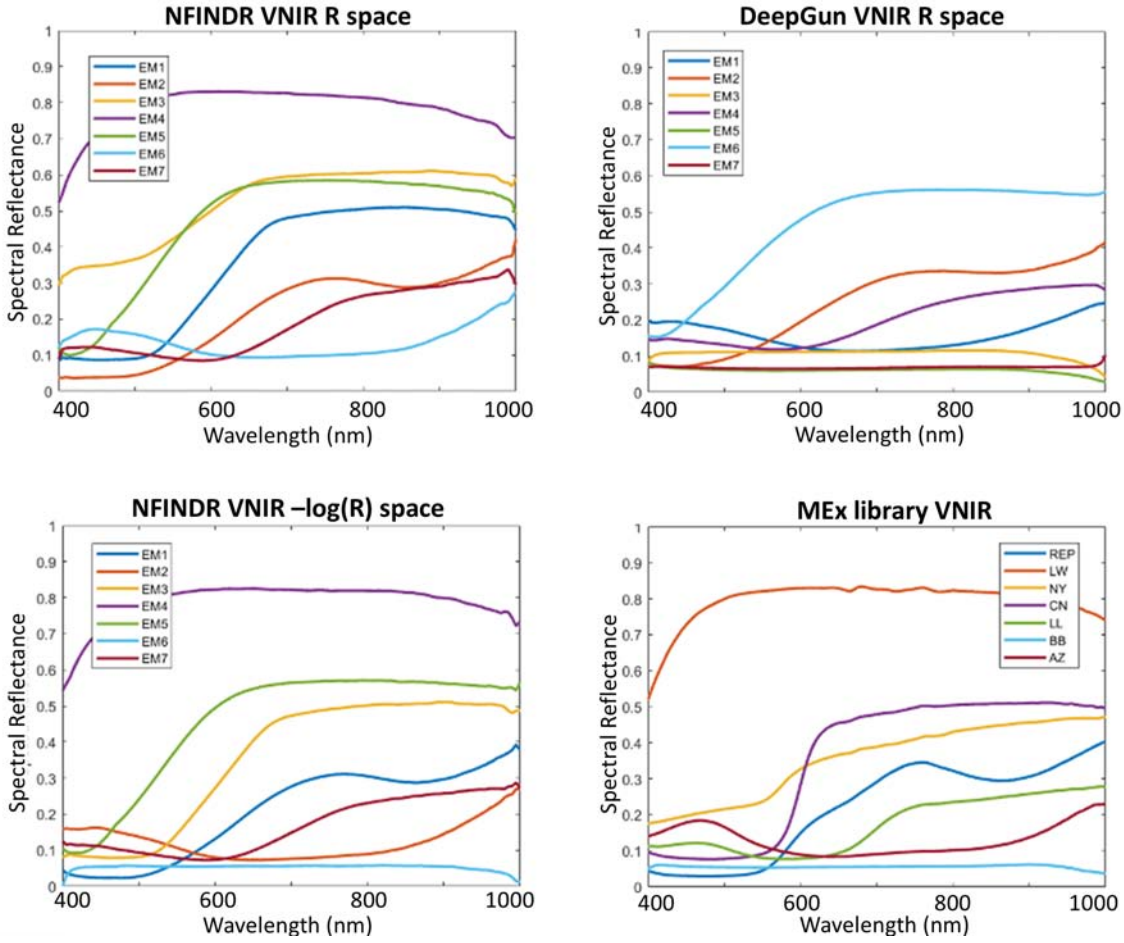

Figure 6.12: EM libraries extracted from the checkerboard reference image in VNIR range. (Upper row) R hyperspace NFINDR (left) and DeepGun (right). (Lower row) -log(R) hyperspace NFINDR (left) and R space MEx (right).

tra for each pixel and the reconstructed spectra (obtained from the estimated concentration vectors and the EM library using the mixing model) is used as part of the evaluation procedure.

In Table 6.3, the $cGFC$, $RMSE$ and ΔE_{00} values obtained from the comparison between the estimated spectra and the original spectra are shown for each EM library, range, and hyperspace. In -log(R) hyperspace, the estimated concentration vectors were obtained using the additive mixture model, but the estimated spectra were obtained directly using the subtractive model with the concentration vectors

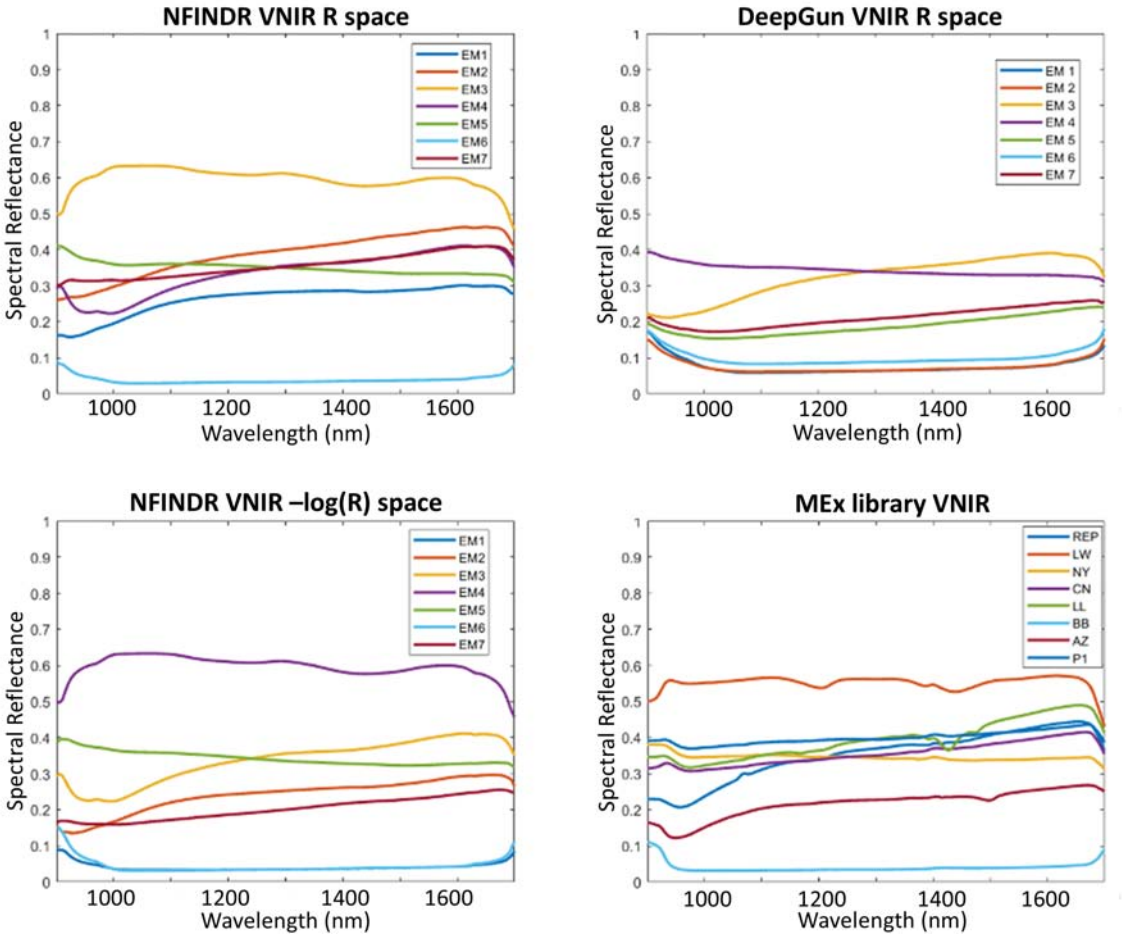

Figure 6.13: EM libraries extracted from the checkerboard reference image in SWIR range. (Upper row) R hyperspace NFINDR (left) and DeepGun (right). (Lower row) -log R hyperspace NFINDR (left) and R hyperspace MEx (right).

and the EM libraries.

Regarding the results in R hyperspace, the best estimated spectra are obtained with the MEx library, as expected. The second-best results correspond globally to NFDL, which is close to DeepGun results in most metrics.

However, in -log(R) hyperspace, the best results correspond to NFDL in all metrics (VNIR range), and in $cGFC$ in the SWIR range. The best results overall correspond to NFDL in -log(R) hyperspace in both ranges, since the slight difference in $RMSE$ in favor of the MEx library does not compensate for the huge difference in

Table 6.3: Spectral reconstruction quality metrics for the three EM libraries in VNIR and SWIR, and in both hyperspaces. The best results for each metric and spectral range are in bold.

Library	Hyperspace	Mean cGFC (STD)	Mean RMSE (STD)	Mean ΔE_{00}	Range
NFDL	R	0.0070 (0.0122)	0.0425 (0.0388)	7.03 (5.00)	VNIR
DeepGun		0.0057 (0.0075)	0.0655 (0.090)	7.70 (5.92)	
MEx		**0.0016 (0.0017)**	**0.0163 (0.0128)**	**5.11 (4.78)**	
MEx	-log(R)	0.0018 (0.0020)	0.017 (0.015)	4.91 (5.08)	
NFDL		**0.0016 (0.0032)**	**0.0098 (0.0060)**	**1.53 (1.61)**	
NFDL	R	0.0021 (0.0040)	0.0172 (0.0105)	-	SWIR
DeepGun		0.00065 (0.00067)	0.0387 (0.0562)	-	
MEx		**0.00068 (0.0015)**	**0.0064 (0.0033)**	-	
MEx	-log(R)	0.00078 (0.0018)	**0.0057 (0.0027)**	-	
NFDL		**0.00027 (0.00034)**	0.0061 (0.0047)	-	

$cGFC$ values in the SWIR range.

In general, the -log(R) transformation is very beneficial for the NFDL in both ranges, but for MEx the quality improves only for ΔE_{00} in VNIR and $RMSE$ in SWIR (both are scale-sensitive metrics). Since MEx is exactly the same in both hyperspaces, this means that performing the optimization in the -log(R) hyperspace with an additive model is not necessarily providing better performance in all cases (i.e., there is not an intrinsic advantage of changing hyperspace). However, the noticeable improvement in NFDL quality points out that it is indeed beneficial to ensure a correspondence between the mixing models in the concentration vector estimation and the EM extraction parts of the unmixing.

VNIR and SWIR results can be compared for $cGFC$ and $RMSE$. The best results correspond to the SWIR range in both cases, with an average across conditions almost three times less than the VNIR range for $cGFC$, and around 20% less for $RMSE$. This means that in general, the estimation of the image reflectances can be more accurate in the SWIR range. This is expected because the SWIR reflectances tend to be flatter and with less variation in range than the VNIR reflectances (see Figure 6.12).

Considering acceptable accuracy values below 1 ΔE_{00}, 0.02 $RMSE$ and 0.005 $cGFC$ (tentative values based on previous experience with spectral estimation), then even the best results would be acceptable in most cases for $cGFC$, while only for MEx in R hyperspace in the VNIR range for $RMSE$, and in all cases in -log(R) hyperspace. None of the results would be acceptable for ΔE_{00}, although NFINDR in -log(R) hyperspace comes close to the threshold. This is expected because the optimization was carried out using $RMSE$ and $cGFC$ values, but not ΔE_{00}.

Concentration and presence maps

The concentration maps show, in grayscale, the concentration of the EMs for each pixel of the checkerboard reference image, ranging between 0 (black) and 1 (white). In Figure 6.14, the concentration maps corresponding to the three libraries tested in the VNIR range and R hyperspace are shown, along with the labeled checkerboard reference image. For NFDL and DeepGun, the EMs are not directly mapped into pigments, as it happens for the MEx library. But sometimes correspondences can be induced from the concentration maps' results. For instance, if an EM concentration map shows significant presence of the EM in patches 5 (LL) and 8 (LL+LW), with higher concentration values in patch 5, and just negligible traces in the other patches, it is safe to assume that this EM corresponds to the LL pigment.

Even for the MEx library, the results of the unmixing as judged by the concentration maps shown in Figure 6.14 (right column) are not completely satisfactory. For instance, EM3 (NY) is not detected with a high concentration in patch four, and the two red pigments (REP and CN, EM1 and EM4) tend to be confused to a certain extent. On the other hand, LW is correctly identified as present in the three mixed patches (8,9 and 10) in EM2 concentration map. The NFDL concentration maps (Figure 6.14, left column) present certain similarities with the MEx results. For instance,

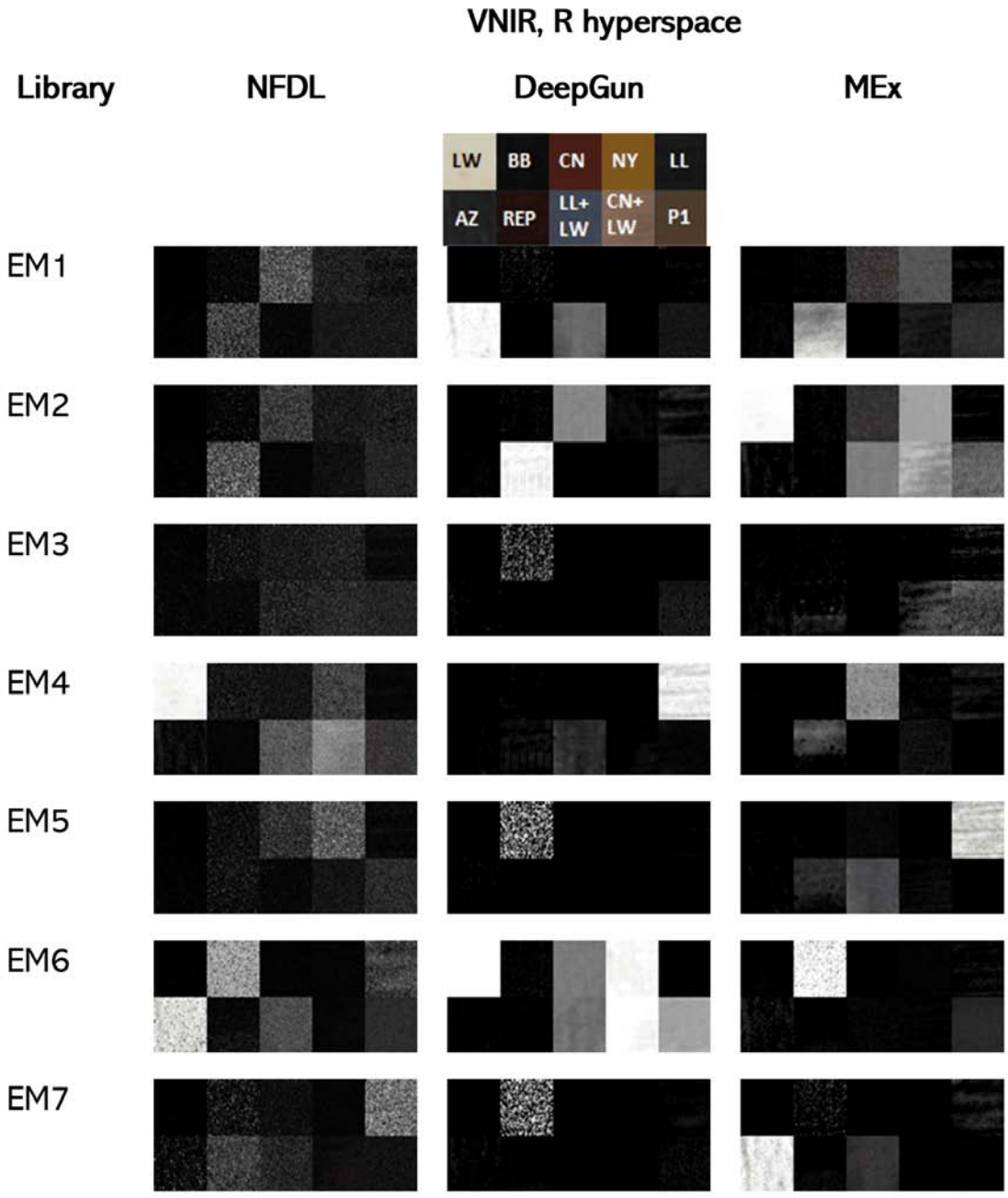

Figure 6.14: VNIR range concentration maps in R hyperspace for the checker reference image and the three EM libraries tested.

EM4 concentration map in NFDL is rather similar to EM2 (LW) for manual extraction, and EM1 from NFDL is similar to EM4 (CN) of the MEx library. The DeepGun concentration maps (Figure 6.14, middle column) show relatively good results

for EM4 (presumably corresponding to LL), EM6 (which has similar appearance to the LW, MEx EM2, concentration map) and EM2 (presumably REP). However, EM1 (presumably AZ) is confused with the LL present in patch 8, and the other three concentration maps (EM3, EM5 and EM7) are very much alike.

In Figure 6.15, the SWIR range R hyperspace concentration maps are shown. These reflect less similarity between NFINDR and MEx libraries results, although with some exceptions (see EM6 in the left and right columns of Figure 6.15). In general, none of the libraries is able to correctly reproduce the real contents of the reference image. There is a trend to confuse again LW with NY and CN for NFDL (see EM3 in Figure 6.15 left), but not for MEx (see EM2 and EM3 in Figure 6.15 right column; however, NY is incorrectly found in patches 8 and 10 in the EM3 concentration map). The CN pigment (patch 3 in the checkerboard image) is not clearly correlated with any of the NFDL EMs, but for MEx (see EM4, Figure 6.15 right) it tends to be confused with LL. The DeepGun concentration maps are more varied in this range, although with similar problems of confusing CN and REP, LW and NY, and AZ with LL, or finding LL in the P1 mixture (see EM7 map in Figure 6.15 middle column). In this range and hyperspace, overall, the DeepGun results seem to be more satisfactory than the NFDL results.

In Figure 6.16, we show the concentration maps for NFDL in -log(R) hyperspace, along with the MEx concentration maps for comparison. In the this hyperspace, NFDL clearly benefits from the change to an additive model (see Figure 6.16 left), being able to correctly identify the presence of REP (EM1) in patches 7 and 10; AZ (EM2) in patch 6 (although this pigment has a trend to be confused with BB); CN (EM3) in patches 3 and 9 (with a trend to confuse REP with CN as well), LW (EM4) in patches 1, 8, 9 and 10; NY (EM5) in patch 4 with a slight residue in patch 10; BB (EM6) in patch 2 (with a trend to appear as well in patch 6 which is AZ, but a slight

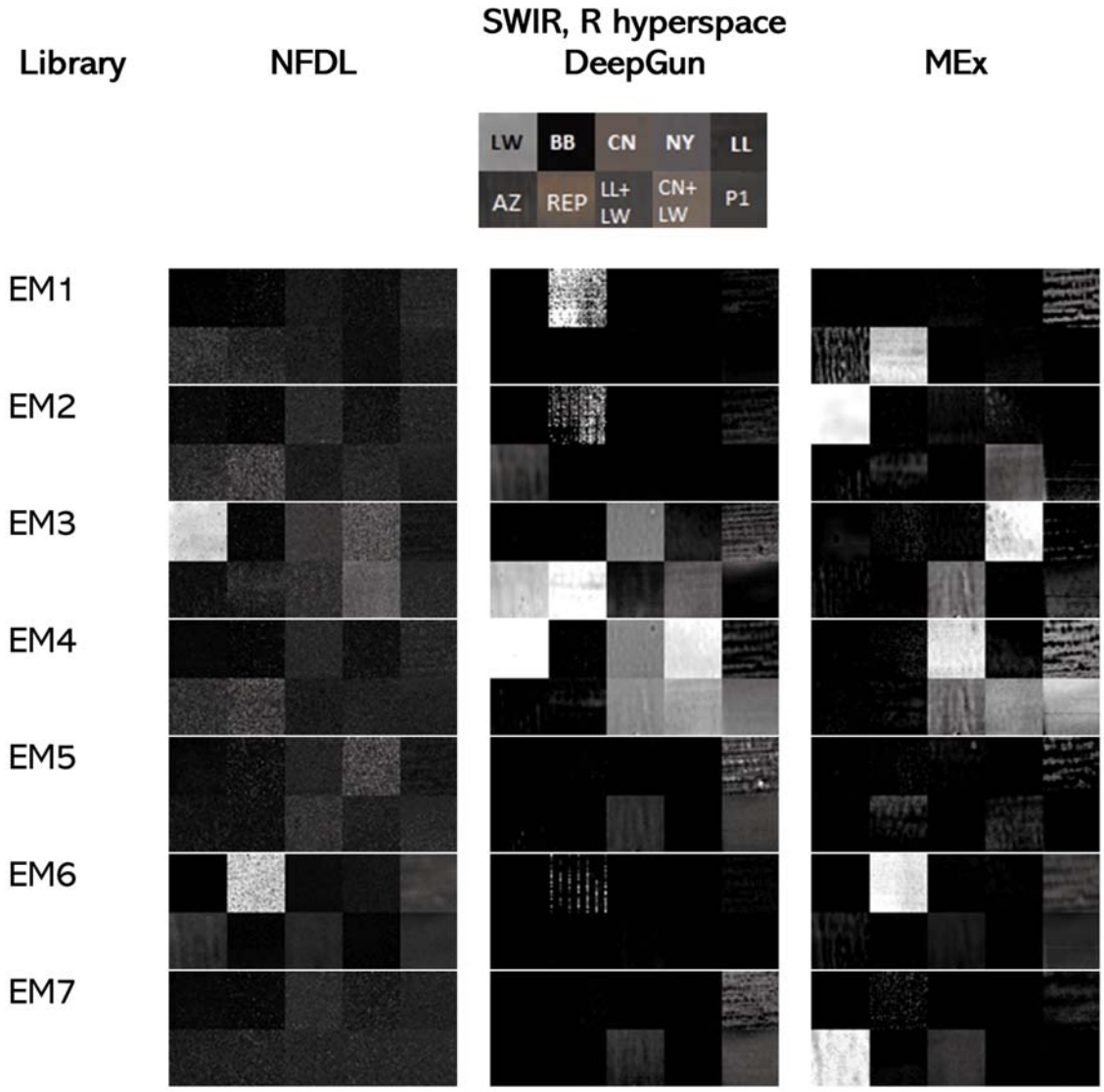

Figure 6.15: SWIR range concentration maps in R hyperspace for the checker reference image and the three EM libraries tested.

concentration value in patch 10 which contains BB), and finally LL in patches 5 and 8, without any confusion with AZ as it happened for the R hyperspace. This is the best and most coherent result, and, in this hyperspace, it is better globally than the results of the MEx library, which retains the confusion between LW and NY (EM2) and REP and NY (EM1). The presence maps corroborate the results obtained with the concentration maps, although some patches with residual traces in the concentration

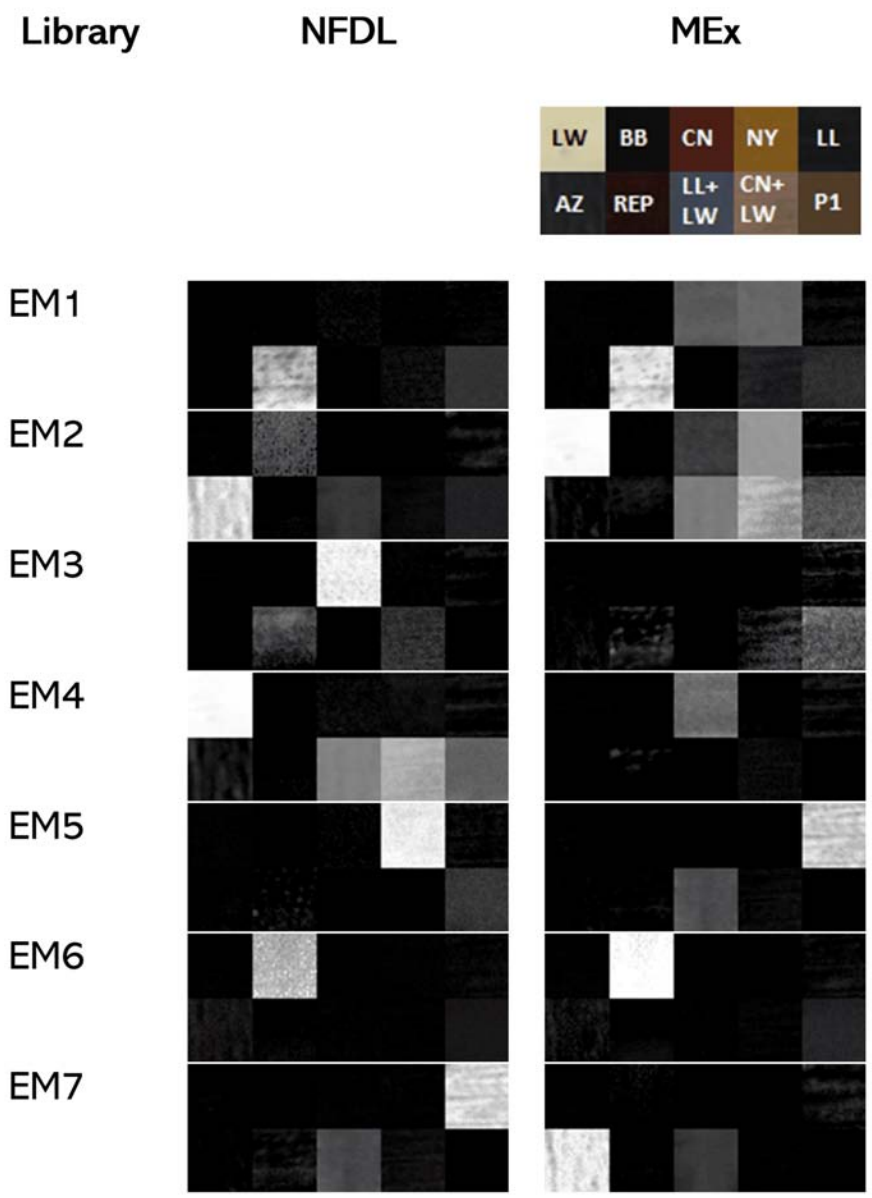

Figure 6.16: VNIR range concentration maps in -log(R) hyperspace for the NFINDR library (left column) and MEx library (right column).

maps are shown in black in the presence map (like EM2 and EM6 in Figure 6.16 left).

Then, in the VNIR range the best model in estimation quality and concentration and presence maps is NFDL in -log(R) hyperspace, with the drawback of a trend to confuse BB and AZ and REP with CN.

Table 6.4: Pigment identification results for NFINDR and DeepGun libraries, using REFL library as reference.

Library	Hyperspace	Range	Assigned labels	Hit rate	MId range
NFDL	R	VNIR	CN, REP, P1, LW, P1, AZ, LL	71.4	0.026-0.365
	-log(R)		REP, AZ, CN, LW, P1, BB, LL	**85.7**	0.021-0.357
DeepGun	R		AZ, REP, BB, LL, BB, P1, BB	57.14	0.0386-0.3181
NFDL	R	SWIR	AZ, LL, LW, REP, NY, BB, CN	**100**	0.0021-0.0251
	-log(R)		BB, AZ, REP, LW, NY, BB, AZ	71.4	0.0069-0.0251
DeepGun	R		BB, BB, REP, NY, AZ, BB, AZ	57.14	0.0049-0.0488

For the SWIR range, the most consistent results are still produced by NFDL in the -log(R) hyperspace (not shown for brevity). The results are slightly less consistent than for VNIR range, with a trend to identify blue content (LL) in red patches (CN and REP and its mixtures), and to identify LW in the CN patch. The MEx library presents several problems as well (finding CN in LL and P1 patches, REP in LL and AZ patches, and NY in mixtures of LL or REP in patches 8 and 10).

Summarizing, even if the concentration maps were not totally satisfactory, consistent results were offered by NFDL in -log(R) hyperspace, especially for the VNIR range. In general, SWIR range tends to offer less consistent results, which might be caused by less varied spectra across pigments.

Pigment identification results

In Table 6.4, the results of the pigment identification method are summarized, including the assigned labels, the hit rate over the seven correct pigments (LW, REP, CN, BB, LL, AZ, NY) and the range of values of MId for the seven pigments. The tolerance value for MId is around 0.03 in VNIR and 0.02 in the SWIR range, given that MId does not include the ΔE_{00} term in the SWIR. The REFL library includes the P1 mixture as well, so it contains eight labels. A perfect pigment identification result would identify only pure pigments from REFL.

The best results are obtained by the NFDL in the SWIR range and R hyperspace,

with a 100% hit rate and maximum MId values very near the tolerance of 0.02 for all EMs. The second-best results correspond to NFDL -log(R) hyperspace in the VNIR range, with a hit rate of 85.7% (six out of seven pigments correctly identified). DeepGun obtains at most four out of seven pigments in both ranges. The P1 mixture is identified as well as part of the pigment palette in all cases for the VNIR range. This does not happen in the SWIR range. Despite the very good results of the unmixing process for -log(R) hyperspace in the VNIR range and NFINDR, the pigment identification is slightly better in the SWIR range and R hyperspace for this library. This suggests that it is convenient to include both ranges in the pigment identification procedure. A control identification procedure performed with MEx library yielded the expected results (hit rate of 100% and MId range very close to 0).

6.3.2.2 Proposed method for analysis of the painting

Given the results for the copper reference plate with known and regular pigment distribution, the following method will be employed for obtaining concentration or presence maps and pigment identification for the painting on copper:

1. Use the -log(R) hyperspace and NFDL_p with 5 EMs extracted from a subsampled spectral image (1:2 ratio) to obtain the concentration and presence maps. The subsampling allows to reduce the computation time required for the unmixing.

2. Use the REFL library extracted from the copper plate to perform pigment identification in both R and -log(R) hyperspaces with NFDL_p.

In this case, we expect lower rates of success in the pigment identification phase, because the painting has been aged for a long period of time and the raw materials

used for the pigments and binders might not be exactly the same in the painting and reference copper plate.

6.3.2.3 Painting on copper

Although we will show the results corresponding to the method proposed in subsubsection 6.3.2.2, a complete analysis using DeGu_p and MEx_p libraries was carried out. The results of these libraries will be commented on only when they outperform NFDL_p.

Unmixing results

Extracted endmember libraries

In Figure 6.17, the NFDL_p libraries in VNIR and SWIR range in -log(R) hyperspace are shown along with the MEx_p library. The NFDL_p EMs are clearly higher in scale than MEx_p EMs, and they tend to be less flat in the SWIR range. The NFDL_p EMs are clearly higher in scale than MEx_p EMs, and they tend to be less flat in the SWIR range.

Spectral reconstruction quality

In Table 6.5, the quality metrics are shown for the NFDL_p (based on the NFINDR algorithm and a linear mixing model), DeGu_p (based on Deep Learning for EM extraction and a non linear model) and MEx_p (based on manual EM extraction) libraries in both spectral hyperspaces and both spectral ranges.

In the VNIR range, the best results (this is, the lowest values of spectral metrics) correspond to MEx_p in -log(R) hyperspace, although the best $cGFC$ value is found in R hyperspace for this library. The metric values found for NFDL_p in -log(R) hyperspace are close to the MEx_p library, with the lowest $RMSE$ values. However, in R hyperspace the NFDL_p estimation is clearly worse than MEx_p estimation,

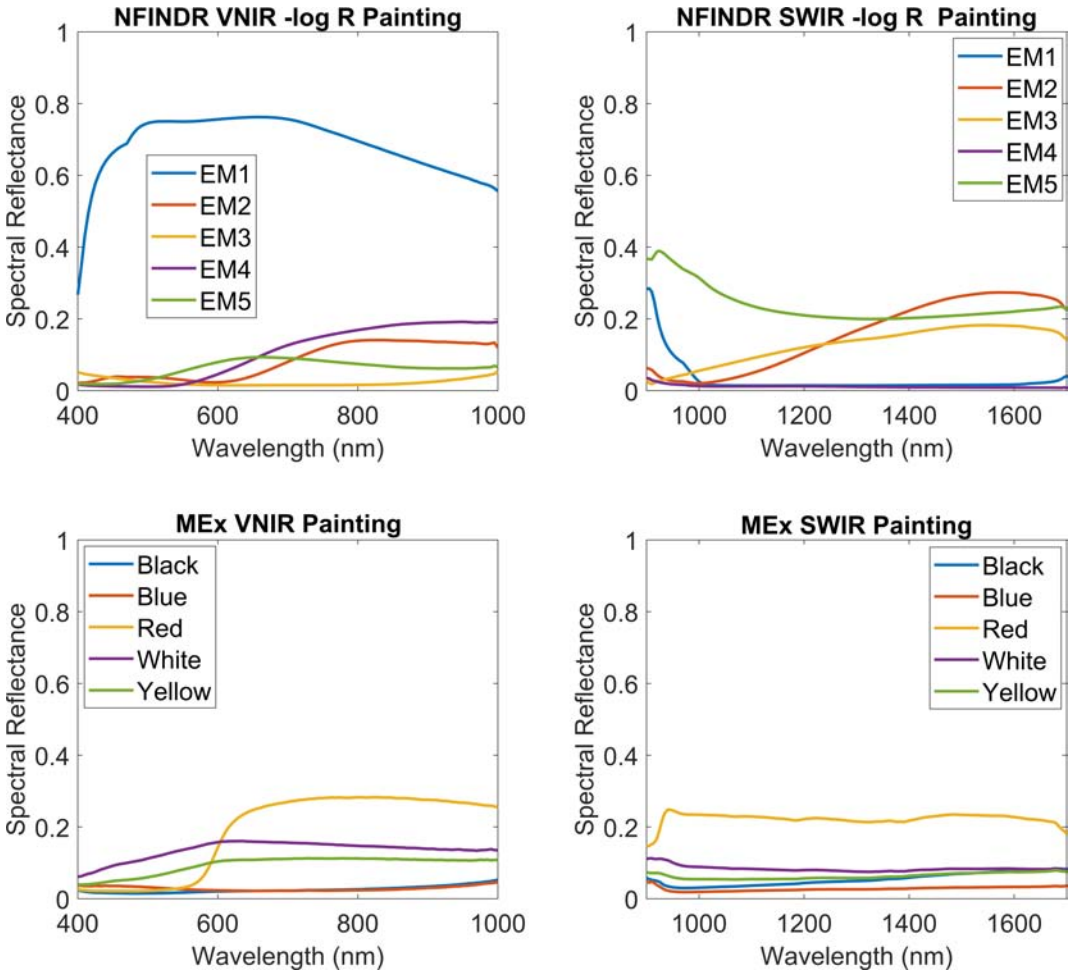

Figure 6.17: VNIR range libraries (left) and SWIR range libraries (right) extracted with NFINDR (upper row) and by Manual Extraction MEx_p (Lower row) from the painting on copper. EMX stands for EM X.

and DeGu_p outperformed NFDL but not MEx_p.

In the SWIR range, the best results for $RMSE$ are found again for NFDL_p in -log(R) hyperspace, and for $cGFC$ in DeGu_p in R space. In general, the quality of the estimation is acceptable or remarkably good, depending on the hyperspace and range, save for NFDL_p in VNIR range and R hyperspace.

The beneficial effect for the scale sensitive metrics ($RMSE$ and ΔE_{00}) of the -log(R) hyperspace transformation is remarkable, and overall, the SWIR estimated spectra are closer to the original according to the spectral metrics, which is expected

Table 6.5: Spectral reconstruction quality metrics for the two EM libraries in VNIR and SWIR, and in both hyperspaces. The best results for each metric and spectral range are in bold.

Library	Hyperspace	Mean cGFC (STD)	Mean RMSE (STD)	Mean ΔE_{00} (STD)	Range
NFDL_p	R	0.0104 (0.010)	0.0743 (0.0309)	12.64 (5.11)	VNIR
MEx_p		**0.0025 (0.0036)**	**0.0136 (0.0146)**	**4.16 (2.29)**	
DeGu_p		0.0037 (0.0036)	0.0324 (0.0270)	6.22 (4.39)	
NFDL_p	-log(R)	0.0065 (0.0095)	**0.0055 (0.0047)**	3.81 (2.48)	
MEx_p		**0.0051 (0.0074)**	0.0076 (0.0125)	**3.69 (2.50)**	
NFDL_p	R	0.0032 (0.0103)	0.0108 (0.0121)	-	SWIR
MEx_p		0.0051 (0.0074)	**0.0076 (0.0125)**	-	
DeGu_p		**0.0016 (0.0030)**	0.0244 (0.0213)	-	
NFDL_p	-log(R)	**0.0035 (0.0055)**	**0.0047 (0.0100)**	-	
MEx_p		0.0090 (0.0267)	0.0064 (0.0010)	-	

because they have lower maximum values.

Presence and concentration maps

In Figure 6.18, the concentration maps for the NFDL_p and MEx_p libraries in -log(R) hyperspace and VNIR range are shown. The order of the MEx_p EMs corresponds to Figure 6.17.

According to the XRF and XRD results [293], LL can be found both in the Virgin's mantle and in the background. This is corroborated by the VNIR range concentration maps EM1 and EM2 for MEx_p (second row), and EM3 for NFDL_p (first row of Figure 6.18). The CN pigment is found in the Virgin's dress and in the carnations, which is corroborated by MEx_p EM3 (second row) and NFDL_p EM4 (first row). The LW is found in the carnations and in the Virgin's chemise sleeves, which is corroborated by MEx_p EM4 (second row) and NFDL_p EM1 (first row). The BB pigment is found in the shadowed areas and in the background, which would be supported by MEx_p EM1 (second row) and NFDL_p EM3 (first row). And finally, the NY pigment is found in some parts of the carnations and in the Child's cloth. This would correspond to EM5 in both libraries. The VNIR results are consistent, with some trend to confuse background with LL pigment in some areas for MEx_p.

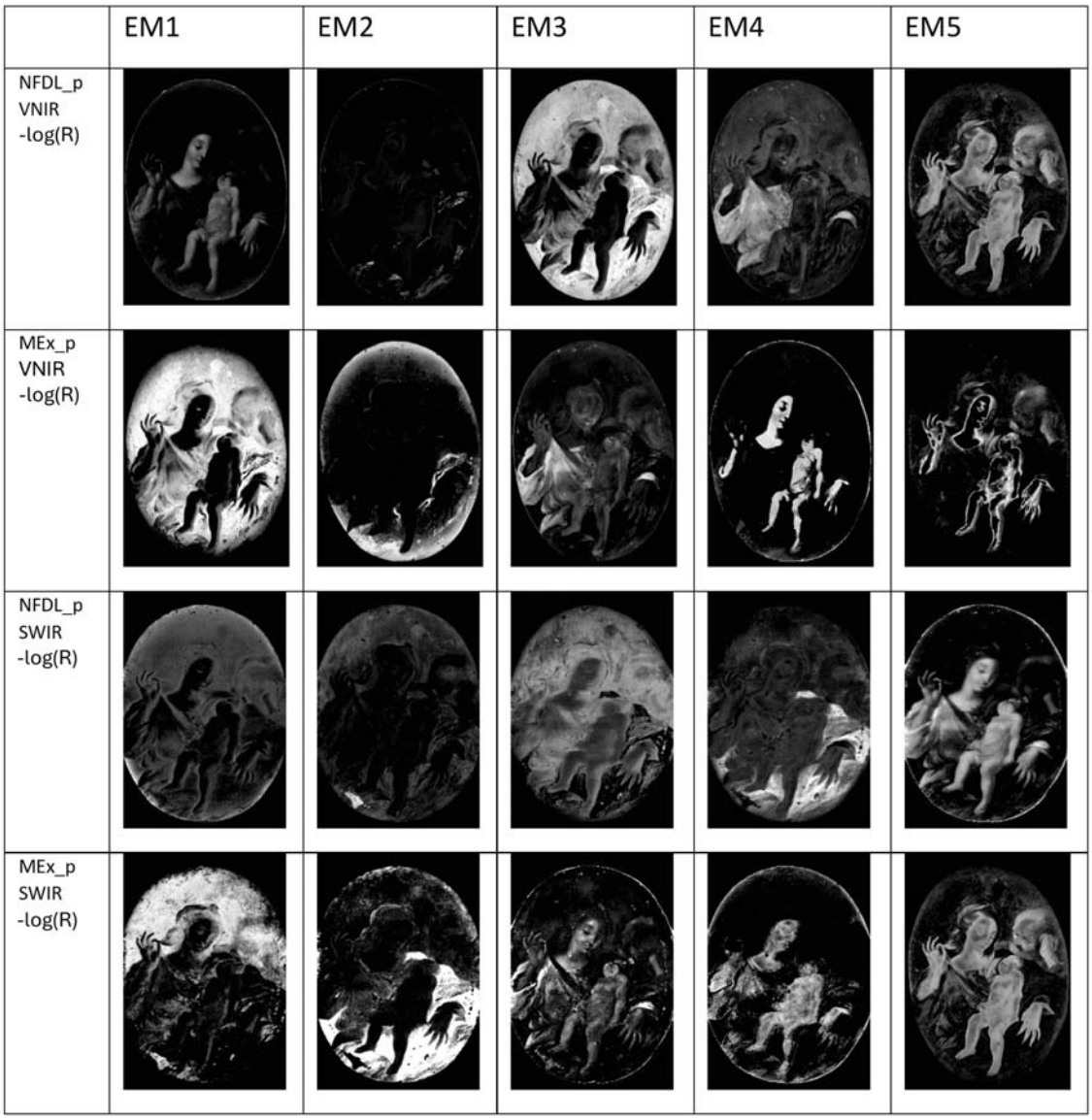

Figure 6.18: Concentration maps in -log(R) hyperspace corresponding to the libraries shown in Figure 6.17 . First row: NFDL_p VNIR range; second row: MEx_p VNIR range; third row: NFDL_p SWIR range; fourth row: MEx_p SWIR range.

In the SWIR range, there are some remarkable findings: the first is that the confusion between BB and LL is less marked for MEx_p (see EM1 and EM2 in the fourth row of Figure 6.18). The second is that for NFDL_p, the LW and CN results are intermingled in EM5 concentration map (third row), while this does not happen for

Table 6.6: Pigment identification results for NFDL_p, using the copper reference REFL library as reference.

Library	Hyperspace	Range	Assigned labels	Hit rate	MId range
NFDL_p	R	VNIR	LL, REP, NY, REP, BB	60	0.13-0.27
	-log(R)		LW, BB, BB, REP, BB	40	0.11-0.41
	R	SWIR	**LW, CN, AZ, NY, BB**	**80**	**0.023-0.15**
	-log(R)		BB, AZ, AZ, BB, NY	40	0.06-0.28

MEx_p (see EM3 and EM4 in the fourth row). And the third is that the NY does not appear clearly in the NFDL_p EMs, while it seems to appear in EM5 for MEx_p (fourth row). The NY pigment is the least present in the painting and appears mostly in mixtures. Overall, the results for the MEx_p library in the SWIR are more consistent, even if the estimation quality is lower than for NFDL_p.

In Figure 6.19, the presence maps with a threshold of 0.25 are shown for both libraries and both ranges, in -log(R) hyperspace. Considering the inherent limitations of the unmixing techniques, the presence map results are rather satisfactory and consistent with the pointwise XRF and XRD results [293]. It is found out that the carnations are a mixture of three EMs, one of them present as well in the Virgin's dress. Or that the Virgin's veil is also a mixture of at least three EMs.

Pigment identification results

In Table 6.6, the results from the pigment identification for NFDL_p using REFL library as reference are shown. The best result is obtained for R hyperspace and SWIR range with a hit rate of 80%, failing to identify the LL pigment. However, if one considers the union of the two ranges, there would be seven pigments, of which only one is not present in the painting (AZ). This union strategy seems to work better also for the -log(R) hyperspace, which would result in six pigments, of which one is not present in the painting (AZ). These results are of course conditioned by the very restricted and specific set of pigments used as reference. Regarding the

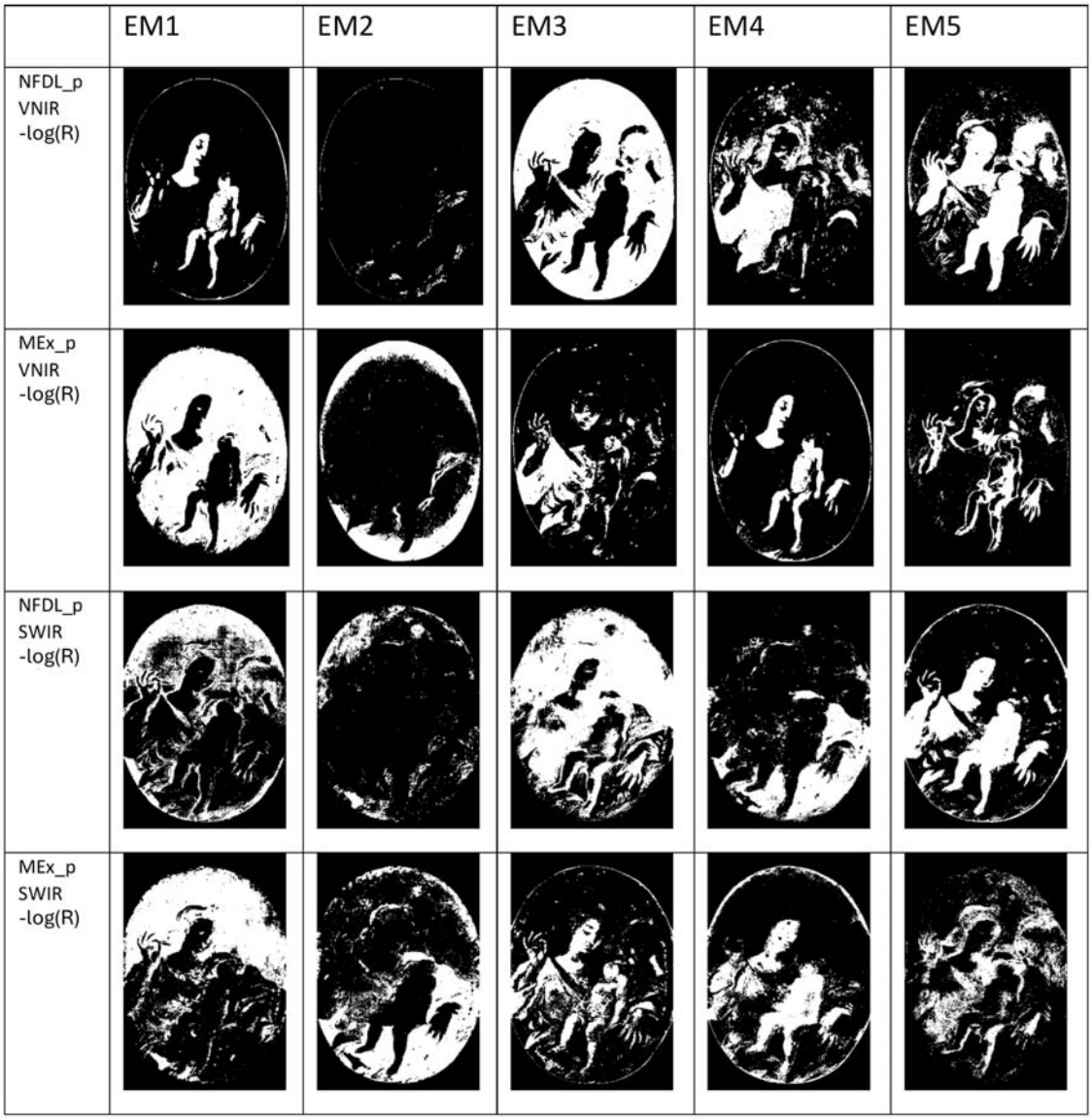

Figure 6.19: Presence maps with a threshold of 0.25, corresponding to the libraries shown in Figure 6.17. Upper row: NFDL_p -log(R) VNIR range; second row: MEx_p -log(R) VNIR range; third row: NFDL_p -log(R) SWIR range; fourth row: MEx_p -log(R) SWIR range.

MId range values, the NFDL_p obtains the lowest value (0.023-0.15) for SWIR and R hyperspace. The results suggest that the SWIR range is more reliable than the VNIR range, in agreement with the hit rate values. None of the other libraries obtain better identification results than NFDL_p.

6.3.3 Conclusions

In this section, different unmixing techniques are applied for pigment detection and identification in two spectral ranges: VNIR and SWIR. The unmixing methods have been selected with the aim to compare two different methodologies: using a non-linear DL-based method in R space, and a linear classical method in -log(R) space.

A reference palette has been prepared, containing the pigments present in an oil painting on copper with the inscription "Boceto di Pablo Veronese" on the back, plus AZ (not present in the painting) and additional patches with mixtures.

The proposed methodology to analyze the painting has been selected using the results obtained with a reference checkerboard image obtained from this palette. For this image, the -log(R) hyperspace unmixing results are satisfactory for the NFDL library in VNIR range in terms of concentration maps and scale dependent quality metrics, although slightly worse for the shape-sensitive $cGFC$ metric. The pigment identification results are also successful.

The method was applied to the restored painting on copper using NFDL_p in both R and -log(R) hyperspaces and in both spectral ranges. NFDL_p was able to provide reasonable results for the concentration and presence maps in the VNIR range. NFDL_p in R hyperspace (SWIR range) was the most successful method for pigment identification. By merging the results of both ranges, the identification would be complete, although AZ was also identified as present in the painting. This supports the usefulness of the SWIR range both for unmixing and pigment identification in artworks.

The results show that, even with unsophisticated techniques (a classical linear algorithm like NFINDR), for some instances it is possible to obtain satisfactory re-

sults using only spectral information for analysis of naturally aged artworks on a somewhat unusual support like copper.

It is crucial to have adequate auxiliary reference pigment palettes for pigment identification using spectral reflectance information, as stated in other studies [134]. The main limitations of the proposed methodology are in pigment identification in the painting, due to the intrinsic differences between the reference pigments (new) and the painting pigments (aged/dirty). Although aging can be modeled to some extent [308] or light-induced pigment degradation can be used [164], multiple variables are involved. The natural aging process changes the pigment's spectral shape differently based on composition, light exposure, and environmental factors. In many cases, these specific factors are unknown, adding complexity to the identification task.

It is worth spending additional effort in refining spectral imaging and unmixing methods so that they can perform on par with alternative techniques without requiring a priori information from the artwork. One limitation is the influence of the preparation techniques, binding agents, varnish, and aging on the reflectance spectra. A multivariate model accounting for the changes introduced by all these factors could lead to more accurate results for stand-alone spectral information-based pigment identification in the future.

CHAPTER 7

Spectral unmixing for historical ink analysis

This chapter presents preliminary results on the application of spectral unmixing techniques for historical ink classification, explored through two different approaches. The first focuses on identifying individual ink components within mixtures of different inks. The second applies spectral unmixing as a pre-processing step to minimize the influence of the support material on the ink spectra, aiming to improve classification performance in an SVM model.

Some of these findings were presented at the *12th Colour and Visual Computing Symposium* (CVCS 2024) and the *XIV Reunión Nacional de Óptica* (RNO 2024), and were further developed during a research stay at the Colourlab, Department of Computer Science at the Norwegian University of Science and Technology (NTNU), Gjøvik, under the supervision of Prof. Sony George.

7.1 Spectral unmixing to identify historical ink mixtures

7.1.1 Introduction

The previous chapter on ML methods for material classification (chapter 5) demonstrated the effectiveness of these algorithms in distinguishing between pure metallo-gallate (MGP), carbon-containing (CC), and non-carbon-containing (NCC) inks in both mock-ups and historical documents, using only HSI data. However, these classifications grouped together a heterogeneous variety of ink formulations. For instance, pure carbon-based inks and their mixtures with metallo-gallate ink or sepia were categorized under the CC label, while mixtures of metallo-gallate ink

with sepia were classified under the NCC label, along with the pure sepia ink (see Table 5.1 for reference).

Medieval Arabic ink recipes suggest that mixed inks played a significant role in the Islamic world [35, 36]. Iron gall ink often exhibited a brownish hue when freshly applied, leading to the practice of mixing it with carbon-based or sepia inks to enhance their color and durability. These mixtures have been documented since ancient times [13, 34]. Despite their historical significance, the detection and classification of mixed inks remain challenging, especially when using non-invasive techniques, and they have received relatively little attention from researchers.

Given the success of spectral unmixing techniques in identifying mixed pigments in a painting (see chapter 6), we extended these methods to the analysis of ink mixtures. The objective was to develop a more accurate and refined classification system than that achieved by ML and DL algorithms, allowing for the mapping and identification of individual ink components from HSI data. Unlike pigments, which exhibit distinct colors and spectral signatures in the VIS range, most historical inks appear black or brown, complicating their differentiation. To address this, spectral analysis was conducted using the concatenated dataset from the VNIR and SWIR ranges through low-level data fusion (as explained in subsection 5.2.3).

Some of the samples analyzed also included pencil marks and various writing supports, which were incorporated into the model. In this way, the method was used to generate concentration maps (for ink identification) and error maps (to evaluate spectral reconstruction) for eight components: three inks (metallo-gallate, sepia, and carbon-based), pencil, and four types of writing supports (parchment, cotton-linen paper, linen paper, and cotton paper). The final result is an unmixing-based classifier not only for inks, but also for pencil and different supports.

7.1.2 Materials and methods

7.1.2.1 Library of endmembers

Spectral unmixing requires a reference library of EM reflectance spectra to determine the concentration of each component pixelwise. Reference spectra were obtained from samples in the *'Mock-ups of historical inks on different supports'* subset (see chapter 3). Specifically, pure samples of metallo-gallate, sepia, and carbon-based inks (including various formulations) applied on parchment and cotton-linen paper were selected, as these were among the most commonly used historical supports present in our database.

Additional reference samples were obtained from historical documents held at the Archive of the Royal Chancellery of Granada and the Provincial Historical Archive, included in the *'Historical documents'* set described in chapter 3. Reflectance spectra were extracted from the hyperspectral images by averaging the pixels corresponding to each class: pure inks, pencil, and support materials, resulting in a reference library with eight classes (EMs): three inks, pencil (Figure 7.1 left), and four writing supports (Figure 7.1 right).

7.1.2.2 Test samples

Spectral unmixing was applied to six mock-up samples from the *'Mock-ups of historical inks on different supports'* subset described in chapter 3. These samples were created by mixing metallo-gallate, sepia, and bone black inks in 1:1 ratios and applying them to parchment and cotton-linen paper. To ensure that results extended beyond controlled mock-ups, spectral unmixing was also applied to a historical manuscript page (referred to as 'parent cube' in chapter 3) from the *'Family tree book from the Archive of the Royal Chancellery of Granada'* subset. This document, also used

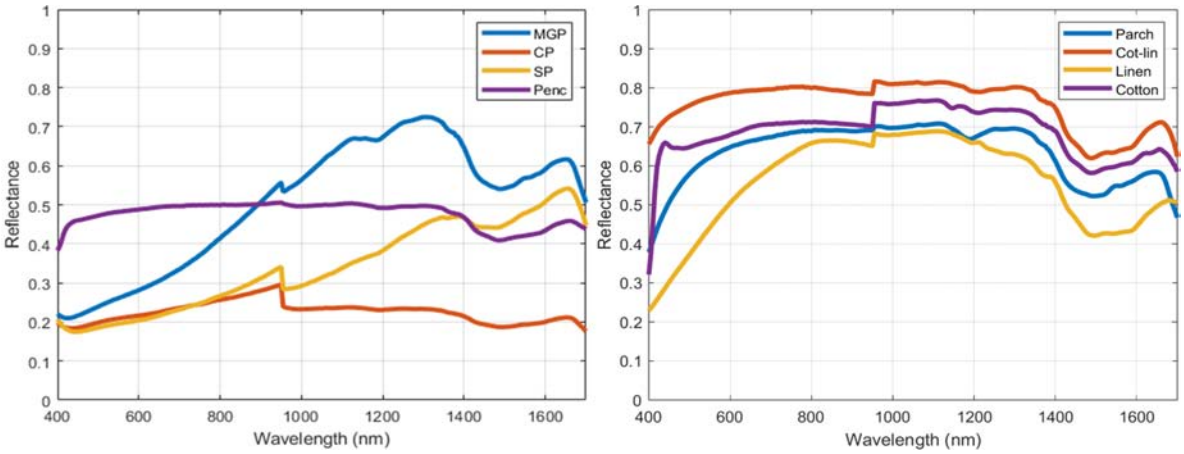

Figure 7.1: Reflectance spectra of the EM library, including foreground materials (left): metallo-gallate pure (MGP), carbon-based ink pure (CP), sepia pure (SP), and pencil (Penc); and background materials (right): parchment (Parch), cotton-linen paper (Cot-lin), linen paper (Linen), and cotton paper (Cotton).

for evaluation in chapter 5, contains handwritten text by two different individuals and includes both carbon-based ink and a mixture of metallo-gallate and sepia inks.

7.1.2.3 Hyperspectral image capture

Hyperspectral images were acquired in the VNIR and SWIR ranges following the capture process and post-processing described in subsection 3.2.2 and subsubsection 3.2.3.1. Data fusion was performed at the low level by concatenating VNIR (400–950 nm) and SWIR (955–1700 nm) spectral information, with interpolation at 5 nm intervals to create a continuous spectral dataset covering the 400–1700 nm range, as described in subsection 5.2.3.

7.1.2.4 Spectral unmixing and ink identification

To determine the concentration of each ink component in the samples, spectral unmixing was carried out using a subtractive mixing model, as proposed by [177]. The optimization algorithm was the interior point [309] with the sum-to-one con-

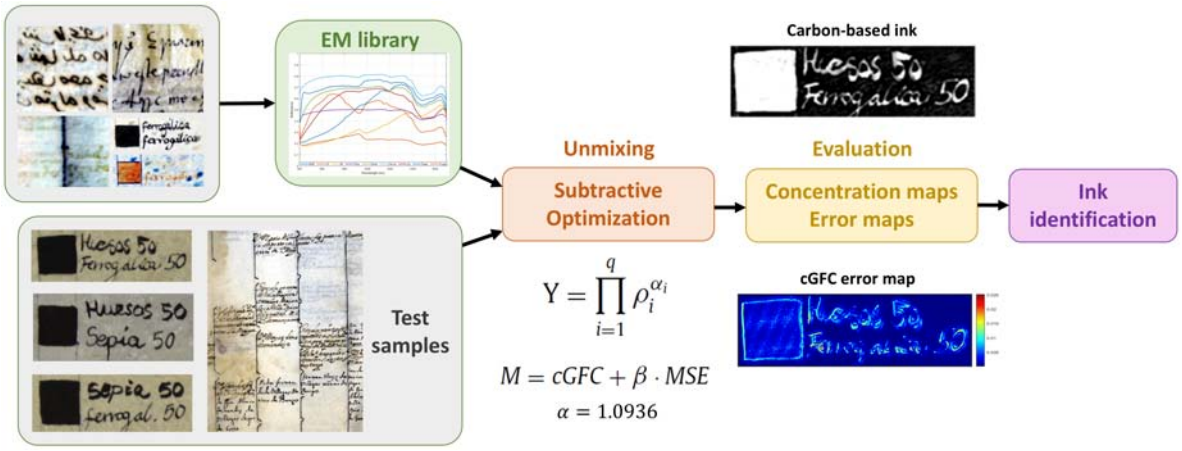

Figure 7.2: Workflow of the proposed methodology for ink unmixing and identification.

straint and a lower bound of zero for the concentrations. An upper bound was not considered, as preliminary tests indicated that it resulted in worse performance. The cost function used incorporated a combined metric of $cGFC$ and $RMSE$ (see subsubsection 6.2.1.4 for details).

The accuracy of the spectral reconstruction was evaluated using error maps ($RMSE$, $cGFC$), and the estimated concentration maps were compared with the known composition of the mock-up samples. For a sample with unknown composition, the EM corresponding to the highest concentration map in the ink/support regions would be selected as the primary component, and the second most concentrated EM would be considered as the second component in the ink mixture. The complete workflow is illustrated in Figure 7.2.

7.1.3 Results

Figure 7.3 (a) presents the concentration maps obtained for the metallo-gallate and bone black ink mixture on parchment. The results show that a higher concentration of carbon-based ink was detected in both the square region and the handwrit-

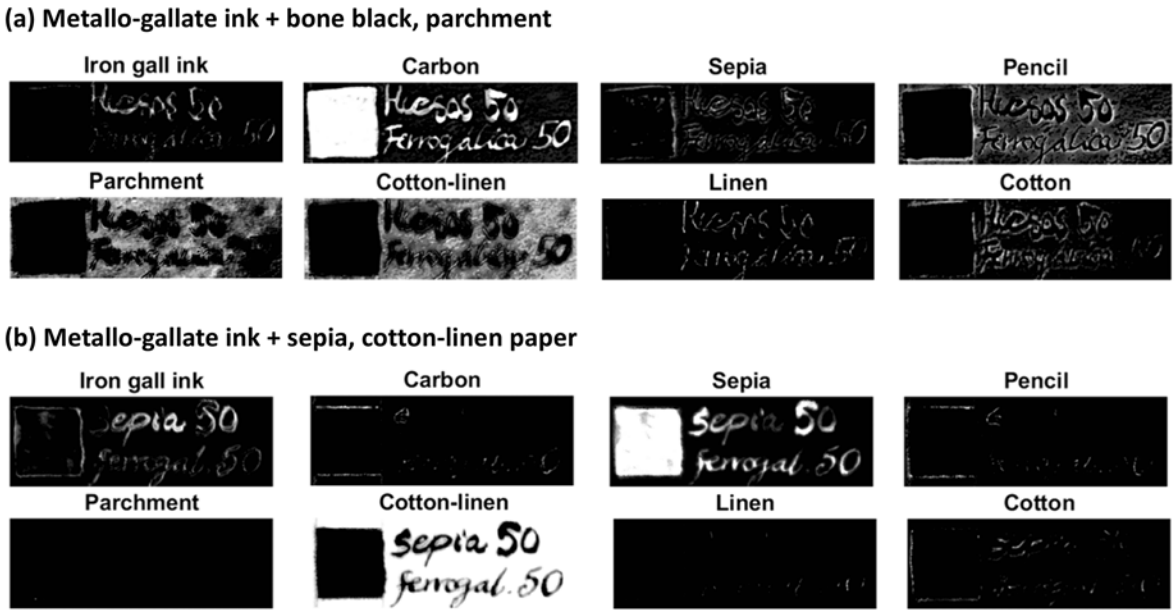

Figure 7.3: Concentration maps of the eight classes in the EM library for mock-ups of (a) metallo-gallate ink and bone black on parchment, and (b) metallo-gallate ink and sepia on cotton-linen paper.

ing traces, while a smaller amount of metallo-gallate ink was identified in the traces. Regarding the supports, differentiation between parchment and cotton-linen paper proved difficult, finding both in similar proportions. The pencil EM was also present on the support, although at a lower concentration. Similar results were obtained for the metallo-gallate and bone black ink mixture on cotton-linen, although the support was correctly identified.

The spectral reconstruction error maps (see Figure 7.4 (a)) indicate higher errors in the square region, where the reconstructed spectrum did not match the original spectrum in amplitude ($RMSE$) but did in overall shape ($cGFC$).

For the sepia and metallo-gallate mixture on cotton-linen paper (see Figure 7.3 (b)), the classification results were satisfactory, as both inks and the support were correctly identified. However, only a small amount of metallo-gallate ink was detected in the square area, and the spectral reconstruction was slightly less accurate

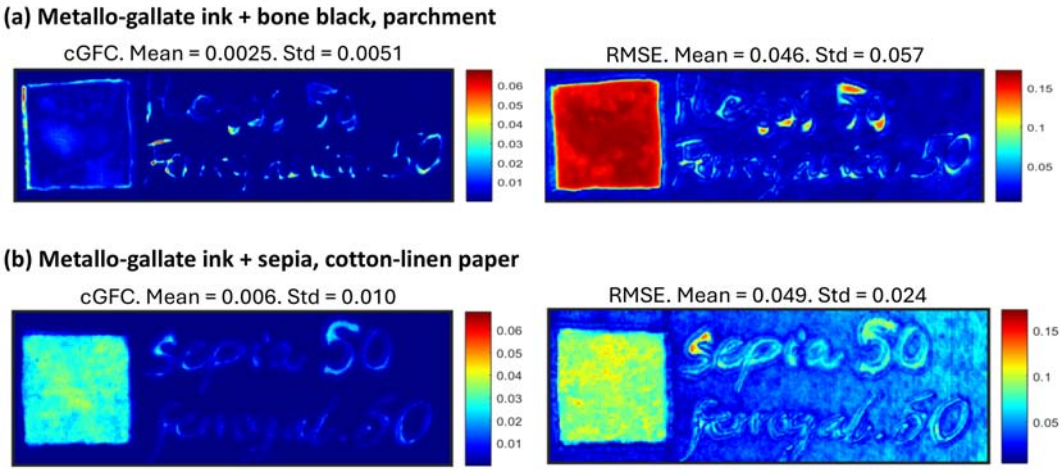

Figure 7.4: Error maps: $cGFC$ (left) and $RMSE$ (right) for mock-ups of (a) metallo-gallate ink and bone black on parchment, and (b) metallo-gallate ink and sepia on cotton-linen paper.

when comparing the mean $cGFC$ and $RMSE$ values in both mock-ups (see Figure 7.4 (b)).

Finally, in the historical document (see Figure 7.5), the paper was correctly identified as cotton-linen, and distinct areas of text were classified as either iron gall ink or carbon-based ink. However, sepia was not detected, and the areas identified as iron gall ink were in fact mixtures of iron gall and sepia.

7.1.4 Discussion and conclusions

As previously discussed (see chapter 3, chapter 4, and chapter 5), metallo-gallate inks become nearly transparent beyond 1000 nm, whereas carbon-based inks continue to absorb infrared radiation. When these two inks are mixed, overall reflectance decreases across all wavelengths, resulting in a spectral profile more closely resembling that of carbon-based ink, making it challenging to detect the presence of metallo-gallate ink even using unmixing techniques.

When comparing the results obtained in this section with those from chapter 5

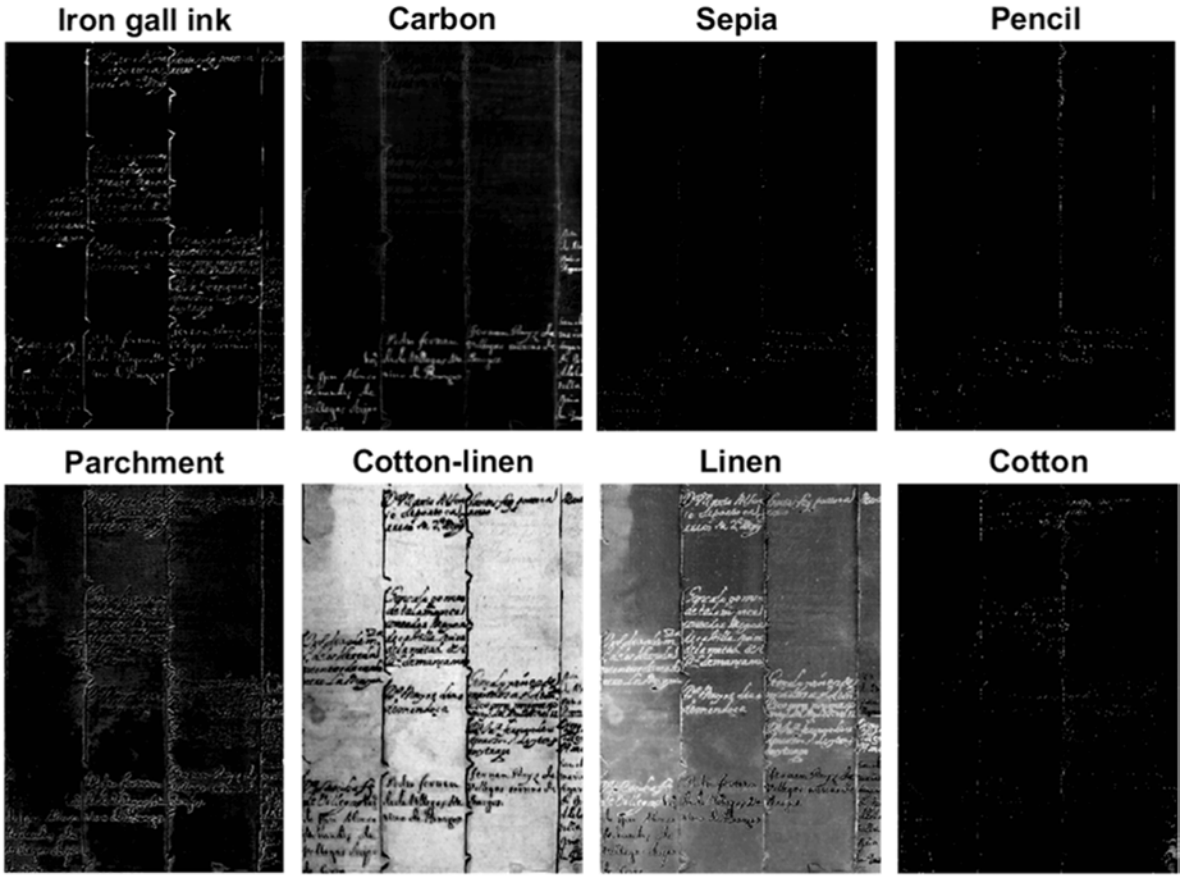

Figure 7.5: Concentration maps of the eight classes in the EM library for a historical document.

for the historical document in Figure 5.9, the NCC and CC classes were detected in different areas. Therefore, the classes created in chapter 5 could serve as a guide for interpreting the unmixing results. In this case, the MGP ink would be excluded due to the absence of the this ink in the classification map of chapter 5, and the pure sepia ink would be discarded because no significant concentration was found in the concentration maps. As a result, the MGP + sepia ink mixture would be the only possible identification.

Further tests are needed to refine the spectral unmixing methodology for ink identification. The presence of carbon-based inks significantly complicates the detection of metallo-gallate and sepia inks. Potential improvements include testing

alternative unmixing models, exploring different cost functions in the optimization, and applying pre-processing steps such as spectral normalization to emphasize spectral shape rather than amplitude. These adjustments could enhance classification accuracy and improve the differentiation of components in mixed inks in historical documents.

7.2 Enhancing ink classification in historical documents using unmixing pre-processing and SVM

7.2.1 Introduction

As discussed in the previous section, the separation of individual components within ink mixtures for identification purposes remains a complex challenge, and the application of unmixing techniques for this purpose requires further optimization to become a viable approach in the future.

As mentioned throughout this PhD thesis, the reflectance spectrum of an ink is not determined exclusively by its chemical composition but is also influenced by other factors, such as the support on which it is applied. This effect is particularly noticeable in the SWIR region, where radiation can penetrate certain inks. Additionally, the boundary between the inked area and the clean support, as well as the presence of very thin strokes, complicates classification due to the spectral mixing of these two components. Consequently, these areas are often misclassified, for example, as pencil marks, as observed in subsection 3.4.1, Figure 3.6, or as different types of ink, as seen in subsection 4.3.1, Figure 4.5, and subsection 5.3.3, Figure 5.7. Furthermore, modern supports differ from aged ones, as in the case of historical documents. As a result, inks with similar compositions, such as pure metallo-gallate

ink, can exhibit significantly different reflectance spectra when applied to mock-ups compared to those found in historical documents, as well as when used on mock-ups with different supports. Some examples of these spectral variations have been discussed in chapter 3.

To address this issue and improve the classification results presented in chapter 5, unmixing has been implemented as a pre-processing step to minimize the influence of the support on the ink's reflectance spectrum, regardless of its aging state. This approach aims to provide the classifier with what we will refer to as 'clean' spectra (i.e., spectra with minimal support influence), potentially facilitating classification independently of the support used. To achieve this, we have to separate the spectral contributions of each component (ink and support) in order to isolate the ink spectrum. An SVM model has been trained and tested (selected as the one providing the best results as traditional ML model in chapter 5), and two new classes, parchment and paper, were introduced into the classifier, resulting in a total of five classes: MGP, CC, NCC, parchment, and paper.

7.2.2 Materials and methods

7.2.2.1 Train/test samples and Machine Learning model

The same training and test samples presented in chapter 5, subsection 5.2.4, were used in this section, with an identical train/test split. However, instead of applying the data fusion approach described in subsection 5.2.3, the logistic splicing correction method [310] was implemented, developed at the Colourlab, where the research stay was conducted. This method effectively mitigates the spectral discontinuity observed between 950 and 955 nm, which arises due to discrepancies in reflectance measurements between the VNIR and SWIR sensors, as previously explained in sub-

section 3.4.2. By applying this correction, the two spectral regions were smoothly concatenated.

A Support Vector Machine (SVM) model was trained and tested (hyperparameters detailed in Table 5.3), as it yielded the best results in the previous study (see chapter 5). In this case, the classes MGP, CC, and NCC were used, along with parchment and paper, to develop a classifier capable of distinguishing between ink and support.

7.2.2.2 Unmixing pre-processing

An automatic unmixing pre-processing approach was developed. For each training sample or minicube (as defined in chapter 3), two EMs were extracted:

1. Ink EM (EM1): This reflectance spectrum was obtained using the GT with the *imerode* function implemented in MATLAB® which selects pixels from the central area of the ink strokes (white region in Figure 7.6 a.). A square structuring element of 5 pixels in width was used for border removal, or 3 pixels for minicubes with very thin strokes. The spectra of the selected pixels were then averaged to form EM1, the ink EM. The resulting spectrum, along with its standard deviation, is shown in Figure 7.6 b. This approach ensured that edge pixels were excluded from the averaging process.

2. Support EM (EM2): A 10×10 pixel square was automatically identified in the GT, located as far as possible from any inked area (red square in Figure 7.6 c.). If a 10×10 square could not be found, a 4×25 pixel rectangle was used instead. The spectra of the pixels within this area were averaged to form EM2, the support EM. The resulting spectrum, along with its standard deviation, is shown in Figure 7.6 d.

These two EMs were then used to generate concentration maps following the procedure described in the previous section and in subsubsection 6.2.1.4, using a subtractive model. Additionally, spectral reconstruction was evaluated using $cGFC$ and $RMSE$.

After this step, two thresholds were selected through various tests to generate presence maps: only pixels with an ink concentration greater than 50% and a support concentration greater than 90% were kept. This was done to exclude pixels at the edges of strokes, where it is difficult to assign a clear label as either ink or support due to spectral mixing.

Following this, an SVM model was trained. Pixels outside the selected thresholds were excluded from training to avoid using data with a high degree of spectral mixing between ink and support. Three different models were then trained based on different spectral processing approaches to the spectral data used for training:

- Model 1: No cleaning – This model used the raw reflectance spectra without any pre-processing, as in chapter 5, but only considering those within the presence maps.

- Model 2: Reconstructed or mixed spectra – In this approach, the spectra used were those reconstructed as the product of the EMs raised to their respective concentrations (subtractive model) after the pre-processing unmixing step.

- Model 3: Clean or unmixed spectra – Ink and support spectra were separated, meaning each was obtained as its corresponding EM raised to its concentration after the pre-processing unmixing step. Unlike in Model 2, no product of the two EM contributions was performed; instead, they were kept separate. The final spectrum retained either the ink or support component, depending on which one had the highest concentration.

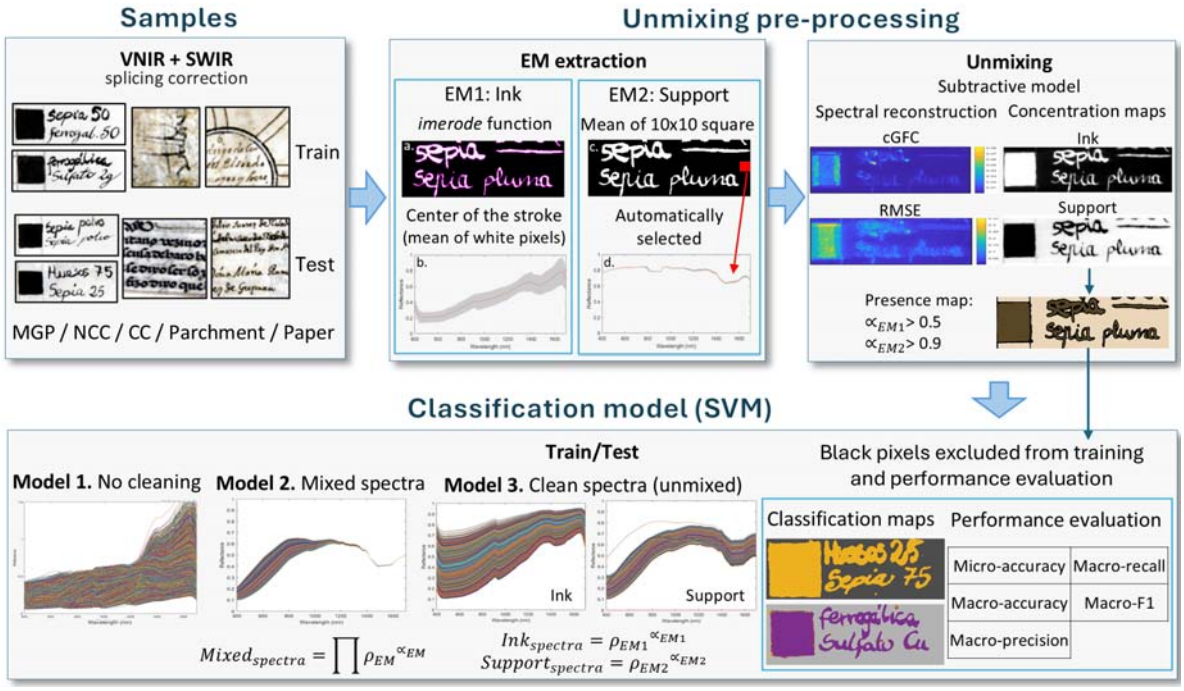

Figure 7.6: Workflow illustrating the steps involved in the process.

Once the spectra were prepared, they were shuffled before training to prevent learning order bias and enhance generalization of the model.

After training, the models were tested. The reflectance values of the pixels in the test samples underwent the same automatic unmixing pre-processing. These reflectance values were then input into each model: raw spectra without pre-processing (Model 1), reconstructed spectra after unmixing (Model 2), and individually reconstructed ink and support spectra (Model 3). The classification results were then used to generate classification maps.

Additionally, the unmixing pre-processing was used to determine the presence maps, as before (with thresholds of 0.5 for ink and 0.9 for support). The evaluation metrics (described in subsection 5.2.8) were applied only to the pixels within the presence maps. This ensured a fair evaluation, as pixels outside these maps contained mixed materials, making it difficult to classify them as a single component

when, actually, they contain a mix of two. Assigning a single label in such cases would not be appropriate. This evaluation procedure was applied consistently to all three models to ensure comparability. The entire process is illustrated in Figure 7.6.

7.2.3 Results

Table 7.1: Mean performance metrics for the SVM model with five classes (MGP, CC, NCC, parchment, paper) on the test set, comparing the impact of different spectral processing approaches (Models 1, 2, and 3).

	Micro-accuracy	**Macro-accuracy**	**Macro-precision**	**Macro-recall**	**Macro-F1**
Model 1	97.81	96.12	92.08	96.12	93.74
Model 2	95.53	96.36	95.38	96.36	95.74
Model 3	97.03	98.12	97.61	98.12	97.80

Table 7.2: Mean performance metrics for the SVM model with only ink classes (MGP, CC, NCC) on the test set, comparing the impact of different spectral processing approaches (Models 1, 2, and 3).

	Micro-accuracy	**Macro-accuracy**	**Macro-precision**	**Macro-recall**	**Macro-F1**
Model 1	95.53	94.40	88.05	94.40	90.64
Model 2	97.88	98.13	95.65	98.13	96.80
Model 3	99.17	99.08	98.84	99.08	98.94

Table 7.1 presents the mean performance metrics for the SVM model in the five-class classification task, distinguishing between metallo-gallate pure (MGP), carbon-containing ink (CC), non-carbon-containing ink (NCC), parchment, and paper. Pencil pixels were excluded from the evaluation, as this class was not part of the classification task but was present in some of the minicubes used for testing. The results show the influence of the different spectral processing approaches (Models 1, 2, and 3) on classification performance, considering only the pixels within the presence maps, as described previously.

For comparison with previous work (chapter 5), paper and parchment pixels were excluded from the test results, and the performance metrics were recalculated

(see Table 7.2). However, support classes may still have appeared in the predictions, as ink pixels could have been misclassified as support. This accounts for the slightly worse results obtained for Model 1 compared to those in chapter 5.

In Figure 7.7, examples of classification maps obtained using the three models with different spectral processing approaches are presented, along with the GT images.

7.2.4 Discussion and conclusions

Regarding the performance metrics obtained for the five-class problem, as presented in Table 7.1, it can be observed that all macro-metrics for Model 2 are higher than those for Model 1, while those for Model 3 exceed those for Model 2. Specifically, there is a 4% difference in Macro-F1 when comparing Model 1 to Model 3. This suggests that the use of unmixing pre-processing can improve classification results in ML models. Micro-accuracy is the only metric where Model 1 shows a slightly higher value than the other models. This can be explained by the fact that micro-accuracy does not account for class imbalance, meaning that classes with more representation in the test set, such as support classes, have a greater influence on the final metric. In this case, Model 2 performed worse when classifying the most represented classes. An analysis of the classification maps suggests that this discrepancy may be due to the misclassification of the support material in one of the mock-up samples (see Figure 7.7 first row). Specifically, in the sample containing a 50:50 mixture of iron gall ink and earth pigment, the support was classified as cotton-linen instead of parchment. This may be due to the optical properties of this ink mixture, which becomes transparent in the IR region. As a result, even though the EM was extracted from the central part of the ink stroke to minimize the influence of the support, the spectral information in the IR range was still primarily influenced by the support material.

Figure 7.7: Classification maps obtained using the SVM model, comparing different spectral processing approaches (Models 1, 2, and 3). The Ground Truth (GT) images are shown in the last column, where black pixels indicate non-evaluated areas. Purple: metallo-gallate ink (MGP); yellow: carbon-containing ink (CC); orange: non-carbon-containing ink (NCC); dark gray: parchment; light gray: cotton-linen paper.

In Table 7.2, when the support classes are excluded from the calculation, it is evident that Model 2 outperforms Model 1, and Model 3 performs better than Model 2 across all the evaluation metrics studied. Specifically, Model 3 achieves more than an 8% higher F1-score compared to Model 1, reaching 98.94%. This performance surpasses the macro-F1 obtained with the DL-based model in chapter 5, although it is important to note that these results are not directly comparable, as here the evaluation is performed on presence maps rather than the entire sample.

Model 1 was designed to closely resemble the one in chapter 5, but with adjusted training and test data to make it more comparable to Models 2 and 3 after pre-processing. The differences between Model 1 in this study and the one used in

chapter 5 are as follows: (i) training was conducted using only pixels from areas with high ink or support concentration; (ii) the labels parchment and paper were included, not just the ink labels (MGP, CC, NCC); (iii) performance metrics for the test evaluation were calculated using only areas with high ink or support concentration. Additionally, the results presented here do not include the post-processing cleaning step proposed in chapter 5. We hypothesize that incorporating this post-processing step would likely improve the results for all three models across most of the samples studied. Some samples, such as those in rows 2, 3, 4, and 5 in Figure 7.7, would benefit from post-processing cleaning. However, as shown in chapter 5, Figure 5.7 (first row) and Figure 5.8 (first row), the SVM model still incorrectly identifies two inks even after applying post-processing. Therefore, Models 2 and 3 would likely perform better than Model 1.

Regarding the classification maps (see Figure 7.7), we observe that in general the areas containing ink are cleaner in Model 2. Model 3 tends to classify the edges of the strokes as pure iron gall ink, which could be addressed with cleaning post-processing. Models 2 and 3 were able to classify the support more uniformly than Model 1 in controlled samples. However, Model 3 presented issues with historical samples, where part of the support was misclassified as pure iron gall ink (see Figure 7.7 4th and 5th rows), particularly in areas with low support concentration according to unmixing (black pixels in GTs). This is understandable, as Model 3 is trained with pure spectra, and the spectrum of pure iron gall ink is the most similar to that of the support, given that this ink becomes transparent in the infrared. This issue could potentially be addressed by adjusting the thresholds or implementing adaptive thresholds, rather than fixed ones (e.g., 0.5 for ink and 0.9 for support), depending on the sample type. Regarding the test samples in the 'Manuscripts of the Provincial Historical Archive of Granada' subset, the misclassifications found in

chapter 5, subsection 5.3.4, persist across all three models. Specifically, some MGP samples are classified as CC, probably due to the effects of ink degradation, aging, and discoloration in the reflectance spectra. This represents a particularly challenging case, where further improvements in the classification process are needed, particularly in addressing the impact of such material changes on the spectral signatures.

After reviewing the results, Models 2 and 3 demonstrate improved classification performance compared to Model 1. Based on an analysis of the classification maps, Model 2 would be preferred due to its uniformity in classification. However, when considering only the classification in the central part of the strokes, Model 3 shows superior performance in that region.

Unmixing emerges as a promising technique to be used as a pre-processing step to improve classification before feeding the data into a ML model. This presents an innovative application of the technique. However, further research is needed to establish it as a fully automated method with adaptive thresholds. One potential approach for classification could involve a voting mechanism, where the materials present in each minicube are determined based on the majority vote from each EM (ink or support). Another approach could involve exploring the use of unmixing for binarization, as it effectively separates ink and support in most cases.

Testing inks on non-porous surfaces, such as glass slides or containers, instead of traditional support materials could help isolate their spectral properties and minimize the influence of the support on the extracted EMs. This would be particularly relevant in the IR range, where certain inks become transparent. However, since inks are inherently applied to support materials in real-world scenarios, a more relevant approach would be to analyze and account for these interactions rather than attempting to eliminate them.

CHAPTER 8

Conclusions and future lines of research

In this PhD thesis, hyperspectral image processing has been applied to the identification of materials in historical documents and artworks. A publicly available database was created to provide a solid basis for spectral analysis. Various approaches were then explored, including spectral similarity metrics, Machine Learning (ML) and Deep Learning (DL) models, as well as spectral unmixing techniques, with the aim of optimizing the classification and mapping of inks, pigments, and supports. The main findings of this thesis are summarized as follows:

1. The HYPERDOC database represents a significant contribution to the field of historical document analysis by providing a comprehensive and publicly accessible hyperspectral dataset of historical documents and mock-ups, being the first of its kind. The dataset includes spatially registered spectral information in both the VNIR and SWIR ranges, along with metadata and Ground Truth (GT) images, offering pixelwise material information. This supports diverse applications, including the development of new methods for material identification and mapping, as explored further in this PhD thesis. This database promotes interdisciplinary collaboration between the image processing and restoration-conservation communities, encouraging the adoption of advanced techniques, such as hyperspectral imaging (HSI), in practical applications within archives and museums. The interest from the community is reflected in the response to the database: within just three weeks of its release, it has received notable media attention, including coverage in newspapers [311–313], television news [314], and social media, as well as over 40 downloads. This increasing engagement suggests that the database's impact will continue to grow in the future.

2. Three methods were proposed to select optimal spectral similarity metrics to classify historical inks, and the performance of six spectral metrics ($RMSE$, SAM, SID, $SIDSAM$, $NS3$, $JMSAM$) was evaluated. In method 1, confidence thresholds were established for each metric, with $SIDSAM$ demonstrating the best per-

formance in ink classification. In method 2, SAM was selected as the metric that provided the highest accuracy and F1-score. In method 3, the estimation of the Probability of Spectral Discrimination (PSD) and the Power of Spectral Discrimination ($PWSD$) resulted in the selection of $JMSAM$ and $SIDSAM$. Combining the results from all three methods proved effective in accurately identifying inks in both mock-ups and ancient manuscripts, although some problems related to the borders of the strokes and the selection of classes justified following up this preliminary study with more advanced ML techniques.

3. Six classification models, including five traditional ML methods (SVM, KNN, LDA, RF, and PLS-DA), and one DL-based model, were implemented for ink classification. Three types of ink were classified: metallo-gallate pure (MGP), carbon-containing (CC), and non-carbon-containing (NCC). Results showed that DL achieved the highest classification accuracy, while SVM performed on par with lower computational requirements. Cleaning post-processing improved the results of traditional models, and data fusion of the VNIR and SWIR ranges proved essential for highlighting distinctive material features, such as fading of MGP inks in the SWIR range. In both case studies, using pages from real historical documents, no model achieved perfect results, likely due to degradation, aging, and fading of CC inks, especially in the SWIR range.

4. Three endmember (EM) extraction methods (NFINDR, NMF and manual extraction) were compared in two stages (pre- and post-intervention) of an oil painting on a copper plate in the VNIR range. The reflectance curves of the NMF EMs differed from real pigments. Manual extraction produced the most accurate concentration maps, closely matching the real pigment distributions, and yielded the best spectral reconstruction. The best pigment identification results were obtained with manual extraction (2 of 5) for the unrestored painting, and NFINDR and NMF (3 of 5) for

the restored painting. The information in the concentration maps was insufficient to detect all re-painted areas, and while spectral band optimization improved their visualization to some extent, it did not enable accurate detection of all areas.

5. Using VNIR and SWIR spectral data, two approaches for automatic EM extraction were compared: NFINDR, based on a linear mixture model, and a DL-based nonlinear model. Two spectral hyperspaces were compared too: the spectral reflectance (R hyperspace) and the -log(R) hyperspace. After evaluating the results on a checkerboard image obtained from a reference palette, NFINDR was selected and applied to the restored oil painting on a copper plate. Unmixing in the -log(R) hyperspace using NFINDR yielded better results in both spectral ranges, while NFINDR in the R hyperspace (SWIR range) was the most successful method for pigment identification (4 of 5). Merging the results from both ranges enabled complete pigment identification, supporting the usefulness of the SWIR range for this purpose.

6. Unmixing techniques were applied to identify individual components in ink mixtures using data fusion of VNIR and SWIR spectral ranges. MGP ink was hardly found when mixed with bone black (a CC ink). This ink becomes transparent in the IR, whereas CC inks continue to absorb this radiation. When mixed, these inks exhibit decreased reflectance across all wavelengths, producing a spectral profile resembling that of CC ink, which complicates the detection of MGP. In the sepia and MGP mixture on cotton-linen, classification results were satisfactory, as both inks and the support were correctly identified. However, sepia was not detected in the historical document. Further refinement of these techniques is needed to achieve reliable results.

7. Applying unmixing techniques as a preprocessing step to minimize the influence of the support on the spectral signature of inks improved classification using SVM. Models 2 and 3, which incorporated unmixing, achieved higher performance

metrics than the model without it. Parchment and cotton-linen classification was also performed. Model 3 frequently misclassified stroke edges as MGP ink, an issue that could be addressed with post-processing cleaning. Model 2 provided the cleanest classification maps. Unmixing emerges as a promising pre-processing technique to enhance ML classification, offering cleaner spectral features for ink identification, though further research is required.

This PhD thesis also opens up several interesting research directions to be explored in the future.

- The HYPERDOC database is a dynamic resource that will continue to grow over time, incorporating additional historical samples from different periods and regions, as well as expanding the variety of materials, including more dyes and pigments for the analysis of illuminated manuscripts.

- Spectral reflectance of inks with similar compositions can change due to aging and discoloration, as seen in the manuscripts from the Provincial Historical Archive of Granada. Future studies should aim to develop classification models that account for these spectral shifts, potentially through virtual de-aging, microfading techniques, and expanding the training dataset as the HYPERDOC database grows.

- As for inks, aging affects the spectral properties of pigments differently depending on composition, light exposure, and environmental conditions. The influence of preparation techniques, binding agents, and varnishes on spectral reflectance should also be explored, with the goal of developing multivariate models to improve pigment identification based on spectral data.

- Future work could include the development of more sophisticated segmentation algorithms for detecting repainted areas, as well as using information from

different spectral ranges.

- Spectral unmixing methods require further refinement for ink analysis. This includes exploring alternative unmixing models and cost functions in optimization algorithms, spectral normalization, and adaptive thresholds for detecting areas with high ink or support content. The potential of spectral unmixing for document binarization could also be explored.

- The interaction between ink and support materials is crucial for spectral analysis, as it alters the final spectrum and complicates classification. The physical properties of the support materials, such as texture, porosity, and color, along with the chemical interactions with the ink, can all influence the spectral signatures, potentially causing spectral shifts or masking certain features. Further investigation is needed to better understand these interactions, as this will improve ink identification using HSI.

- The integration of multiple spectral ranges (VNIR and SWIR) has demonstrated its potential for pigment and ink identification. Future research should focus on optimizing multimodal methods to maximize classification accuracy, such as combining HSI with XRF or Raman spectroscopy. Also, using HSI as a preliminary exploratory technique to provide informed guidance in the selecton of key points for measuring with point-based methods shows the potential value of a combined technique approach for material identification.

HSI is non-invasive, provides high spatial and spectral resolution, and requires significantly less time for capture compared to alternative techniques. These advantages emphasize the need for continued refinement of these methods to improve material identification. By addressing these challenges, future research can build

on the findings of this PhD thesis, contributing to the advancement of HSI in the analysis of historical documents and paintings.

CAPÍTULO 8

Conclusiones y futuras líneas de investigación

En esta tesis doctoral se ha aplicado el procesamiento de imágenes hiperespectrales para la identificación de materiales en documentos históricos y obras de arte. Se ha desarrollado la primera base de datos de acceso público, HYPERDOC, de referencia para el análisis espectral de documentos. A continuación, se exploraron diversos métodos, incluyendo la evaluación de métricas de similitud espectral, modelos de aprendizaje automático y aprendizaje profundo, así como técnicas de *unmixing* o desmezclado espectral, con el objetivo de optimizar la clasificación y determinar la distribución espacial de tintas, pigmentos y soportes. Los principales hallazgos de esta tesis se resumen a continuación:

1. La base de datos HYPERDOC constituye el primer conjunto de datos hiperespectrales de acceso público que integra documentos históricos y muestras modelo. Incluye información espectral registrada espacialmente en los rangos visible e infrarrojo cercano (VNIR) e infrarrojo de onda corta (SWIR), junto con metadatos y *Ground Truth* (GT), que ofrecen información sobre los materiales presentes en cada píxel de la imagen. Este recurso facilita el desarrollo de nuevos métodos para identificar y conocer la distribución espacial de los materiales, promoviendo la colaboración interdisciplinar y la adopción de técnicas avanzadas, como la imagen hiperespectral (HSI), en aplicaciones prácticas dentro de archivos y museos. El interés de la comunidad se refleja en la respuesta a la base de datos: en tan solo tres semanas desde su lanzamiento, ha recibido una notable atención mediática, incluyendo cobertura en periódicos [311–313], noticias televisivas [314] y redes sociales, así como más de 40 descargas. Esto sugiere que el impacto de esta base de datos continuará en aumento en el futuro.

2. Se propusieron tres métodos para seleccionar las métricas de similitud espectral óptimas para clasificar tintas históricas, evaluando el rendimiento de seis métricas espectrales ($RMSE$, SAM, SID, $SIDSAM$, $NS3$, $JMSAM$). En el método

1, se establecieron umbrales de confianza para cada métrica, siendo $SIDSAM$ la que mostró mejor rendimiento en la clasificación de tintas. En el método 2, SAM fue seleccionada como la métrica que proporcionó mayor precisión y F1-score. En el método 3, la estimación de la probabilidad de discriminación espectral (PSD) y el poder de discriminación espectral ($PWSD$) resultaron en la selección de $JMSAM$ y $SIDSAM$. Combinar los resultados de los tres métodos demostró ser efectivo para identificar tintas tanto en muestras modelo como en manuscritos antiguos. Sin embargo, debido a problemas relacionados con los bordes de los trazos y la selección de clases, decidimos aplicar técnicas más avanzadas, como las de aprendizaje automático.

3. Se implementaron seis modelos de clasificación para tintas, incluyendo cinco métodos tradicionales de aprendizaje automático (SVM, KNN, LDA, RF y PLS-DA), y un modelo basado en aprendizaje profundo (DL). Se clasificaron tres tipos de tinta: metalogálica pura (MGP), con contenido de carbón (CC) y sin contenido de carbón (NCC). Los mejores resultados se obtuvieron con DL, mientras que SVM proporcionó un rendimiento similar con una menor demanda computacional. Gracias al post-procesamiento, se mejoraron los resultados de los modelos tradicionales, y la fusión de la información en los rangos VNIR y SWIR demostró ser esencial para resaltar características distintivas de los materiales, como el desvanecimiento de las tintas MGP en el rango SWIR. En el análisis de documentos completos, ningún modelo alcanzó resultados óptimos, probablemente debido a la degradación, envejecimiento y desvanecimiento de las tintas CC, especialmente en el rango SWIR.

4. Se compararon tres métodos de extracción de *endmembers* (EM): NFINDR, factorización no negativa de matrices (NMF) y extracción manual, en dos etapas (pre y post-intervención) de una pintura al óleo sobre placa de cobre en el rango VNIR. Las reflectancias espectrales de los EM obtenidos con NMF diferían de los pigmen-

tos reales. La extracción manual produjo los mapas de concentración más precisos, coincidiendo estrechamente con las distribuciones de pigmentos reales, proporcionando la mejor reconstrucción espectral. Los mejores resultados de identificación de pigmentos se obtuvieron con la extracción manual (2 de 5) para la pintura no restaurada, y con NFINDR y NMF (3 de 5) para la pintura restaurada. La información en los mapas de concentración fue insuficiente para detectar todas las áreas repintadas, y aunque la optimización de bandas espectrales mejoró su visualización hasta cierto punto, no permitió una detección precisa de todas las áreas.

5. Utilizando datos espectrales VNIR y SWIR, se compararon dos enfoques para la extracción automática de EM: NFINDR, basado en un modelo de mezcla lineal, y un modelo no lineal basado en DL. También se compararon dos espacios espectrales: el de reflectancia (R) y el espacio -log(R). Después de evaluar los resultados sobre una imagen obtenida a partir de una paleta de referencia, se seleccionó NFINDR y se aplicó a la pintura restaurada sobre placa de cobre. El desmezclado o *unmixing* en el espacio espectral -log(R) usando NFINDR dio mejores resultados en ambos rangos espectrales, mientras que NFINDR en el espacio R (rango SWIR) fue el mejor método para la identificación de pigmentos (4 de 5). La integración de los resultados de ambos rangos permitió identificar completamente los pigmentos, evidenciando la importancia de utilizar ambos rangos espectrales.

6. Se aplicaron técnicas de desmezclado o *unmixing* espectral para identificar componentes individuales en mezclas de tintas mediante la fusión de datos de los rangos espectrales VNIR y SWIR. Se encontraron dificultades para detectar la tinta MGP cuando era mezclada con negro de huesos (tinta CC), ya que la tinta MGP se vuelve transparente en el infrarrojo, mientras que la tinta CC continúa absorbiendo esta radiación. Como resultado, el espectro de la mezcla muestra una disminución en todas las longitudes de onda, generando un perfil similar al de la tinta CC y

dificultando la detección de la MGP. En la mezcla de sepia y MGP sobre papel de algodón y lino, los resultados de clasificación fueron satisfactorios, ya que tanto las tintas como el soporte fueron correctamente identificados. Sin embargo, la sepia no fue detectada en el documento histórico. Es necesario un mayor perfeccionamiento de estas técnicas para obtener resultados más consistentes.

7. La incorporación de técnicas de *unmixing* como preprocesamiento con el fin de minimizar la influencia del soporte en el espectro de las tintas, mejoró la clasificación mediante un modelo SVM. Los modelos 2 y 3, que incorporaron *unmixing*, lograron un rendimiento más alto que el modelo 1, sin *unmixing*. También se clasificaron los soportes pergamino y papel de algodón y lino. El modelo 3 clasificó erróneamente los bordes de los trazos como tinta MGP, un problema que podría abordarse con post-procesamiento. El modelo 2 proporcionó los mapas de clasificación más precisos. El *unmixing* espectral se presenta como una técnica prometedora de preprocesamiento para mejorar la clasificación con algoritmos de aprendizaje automático, al proporcionar espectros menos afectados por el soporte para la identificación de tintas, aunque se necesita más investigación para perfeccionar su aplicación.

Esta tesis doctoral también abre varias líneas de investigación interesantes que pueden explorarse en el futuro.

- La base de datos HYPERDOC es un recurso dinámico que seguirá creciendo con el tiempo, incorporando muestras históricas adicionales de diferentes periodos y regiones, así como ampliando la variedad de materiales, incluyendo colorantes y pigmentos para el análisis de manuscritos iluminados.

- La reflectancia espectral de tintas con composición similar puede cambiar debido al envejecimiento y la decoloración, como se observa en los manuscritos del Archivo Histórico Provincial de Granada. Estudios futuros deben centrarse

en desarrollar modelos de clasificación que tengan en cuenta estos cambios espectrales utilizando técnicas de des-envejecimiento virtual, *microfading* y ampliando el conjunto de datos de entrenamiento a medida que crece la base de datos HYPERDOC.

- El envejecimiento también afecta de manera diferente las propiedades espectrales de los pigmentos según su composición, exposición a la luz y condiciones ambientales. Asimismo, factores como las técnicas de preparación, los aglutinantes y los barnices utilizados deben considerarse en el desarrollo de modelos multivariantes para mejorar la identificación de pigmentos.

- El desarrollo de algoritmos de segmentación más sofisticados para la detección de áreas repintadas o restauradas, así como el uso de información de diferentes rangos espectrales, es una línea de investigación interesante a explorar.

- Las técnicas de *unmixing* deben ser optimizadas para el análisis de tintas. Esto incluye la exploración de modelos alternativos, nuevas funciones de coste en los algoritmos de optimización, la normalización espectral y el uso de umbrales adaptativos para detectar áreas con alto contenido de tinta o soporte. También se podría explorar el potencial del *unmixing* espectral para la binarización de documentos.

- La interacción entre la tinta y los materiales de soporte es crucial para el análisis espectral, ya que altera el espectro final y dificulta la clasificación. Las propiedades físicas de los materiales de soporte, como la textura, la porosidad y el color, junto con las interacciones químicas con la tinta, pueden influir en la reflectancia espectral. Se requiere mayor investigación para comprender mejor estas interacciones y así mejorar la identificación de tintas mediante HSI.

- La integración de múltiples rangos espectrales (VNIR y SWIR) ha demostrado su potencial en la identificación de pigmentos y tintas. Investigaciones futuras deberían optimizar métodos multimodales, combinando HSI con técnicas como fluorescencia de rayos X o Raman, para maximizar la precisión de clasificación. Además, el uso de HSI como técnica exploratoria preliminar para orientar la selección de puntos clave en mediciones puntuales resalta el potencial de un enfoque combinado para la identificación de materiales.

La imagen hiperespectral (HSI) es una técnica no invasiva que ofrece una alta resolución espacial y espectral, y requiere considerablemente menos tiempo de captura en comparación con técnicas alternativas. Estas ventajas subrayan la necesidad de perfeccionar los métodos existentes para mejorar la identificación de materiales. Al abordar estos desafíos, futuras investigaciones podrán apoyarse en los hallazgos presentados en esta tesis, contribuyendo al avance del uso de HSI en el análisis de documentos históricos y pinturas.

Bibliography

[1] UNESCO Institute for Statistics, "UNESCO Institute for Statistics - Glossary," 2024. Accessed on December 31, 2024.

[2] UNESCO, FCS, "Framework for cultural statistics," *United Nations [Online] Available http://www. uis. unesco. org/culture/Pages/framework-cultural-statistics. aspx*, 2009.

[3] UNESCO, "Convention for the safeguarding of the intangible cultural heritage," *[Online] Available https://ich.unesco.org/doc/src/15164-EN.pdf*, 2003.

[4] UNESCO, "Universal declaration on cultural diversity," *[Online] Available https://www.unesco.org/en/legal-affairs/unesco-universal-declaration-cultural-diversity*, 2001.

[5] H. Stovel, *Risk preparedness: a management manual for world cultural heritage*. International Centre for the Study of the Preservation and Restoration of Cultural Property (ICCROM), 1998.

[6] UNESCO, "Convention concerning the protection of the world cultural and natural heritage," *[Online] Available https://whc.unesco.org/archive/convention-en.pdf*, 1972.

[7] M. Strlič, "Heritage science: A future-oriented cross-disciplinary field," *Angewandte Chemie International Edition*, vol. 57, no. 25, pp. 7260–7261, 2018.

[8] European Committee for Standardization, "CEN/TC 346 - Conservation of Cultural Property," tech. rep., European Committee for Standardization, 2002.

[9] R. Radvan, L. Ratoiu, I. M. Cortea, A. Chelmus, L. Angheluta, and D. Marinescu, "Multi-step approach for characterization of artworks based on hyperspectral imaging and complementary techniques," in *2018 11th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 117–122, IEEE, 2018.

[10] M. J. Melo, V. Otero, P. Nabais, N. Teixeira, F. Pina, C. Casanova, S. Fragoso, and S. O. Sequeira, "Iron-gall inks: a review of their degradation mechanisms and conservation treatments," *Heritage Science*, vol. 10, no. 1, p. 145, 2022.

[11] C. Cucci and A. Casini, "Hyperspectral imaging for artworks investigation," in *Data handling in science and technology*, vol. 32, pp. 583–604, Elsevier, 2019.

[12] J. M. Madariaga, "Analytical chemistry in the field of cultural heritage," *Analytical Methods*, vol. 7, no. 12, pp. 4848–4876, 2015.

[13] S. González-García, A. López-Montes, and T. Espejo-Arias, "The use of writing inks in 12th–19th century Arabic manuscripts: A study of the collection of the School of Arabic Studies, Granada (Spain)," in *Science, Technology and Cultural Heritage*, pp. 121–126, CRC Press, 2014.

[14] F. Lucarelli and P. A. Mando, "Recent applications to the study of ancient inks with the florence external-PIXE facility," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 109, pp. 644–652, 1996.

[15] E. O. Omayio, S. Indu, and J. Panda, "Historical manuscript dating: traditional and current trends," *Multimedia Tools and Applications*, vol. 81, no. 22, pp. 31573–31602, 2022.

[16] A. Polak, T. Kelman, P. Murray, S. Marshall, D. J. Stothard, N. Eastaugh, and F. Eastaugh, "Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication," *Journal of Cultural Heritage*, vol. 26, pp. 1–11, 2017.

[17] F. Daniel, A. Mounier, J. Pérez-Arantegui, C. Pardos, N. Prieto-Taboada, S. F.-O. de Vallejuelo, and K. Castro, "Hyperspectral imaging applied to the analysis of goya paintings in the museum of Zaragoza (Spain)," *Microchemical Journal*, vol. 126, pp. 113–120, 2016.

[18] L. Burgio, "Pigments, dyes and inks: their analysis on manuscripts, scrolls and papyri," *Archaeological and Anthropological Sciences*, vol. 13, no. 11, p. 194, 2021.

[19] COSCH, "COST-Action TD 1201: Colour and Space in Cultural Heritage (COSCH)." `https://www.cosch.info/index.html`, 2016. [Online; accessed 10-Mar-2025].

[20] Norwegian University of Science & Technology, "CHANGE: Cultural Hertiage ANalysis for New GEnerations." `https://change-itn.eu/`, 2023. [Online; accessed 10-Mar-2025].

[21] Universitat de València, "ChemiNova." `https://cheminova.eu/`, 2024. [Online; accessed 10-Mar-2025].

[22] Consejería de Cultura y Patrimonio Histórico. Delegación Territorial en Granada, *Alamas Nazaríes. Los autógrafos de los sultanes (1454-1492)*. Archivo de la Real Chancillería de Granada, Junta de Andalucía, 2022.

[23] I. Ciortan, H. Deborah, S. George, and J. Y. Hardeberg, "Color and hyperspectral image segmentation for historical documents," in *2015 Digital Heritage*, vol. 1, pp. 199–206, IEEE, 2015.

[24] E. Kavallieratou and S. Stathis, "Adaptive binarization of historical document images," in *18th International Conference on Pattern Recognition (ICPR'06)*, vol. 3, pp. 742–745, IEEE, 2006.

[25] A. Sulaiman, K. Omar, and M. F. Nasrudin, "Degraded historical document binarization: A review on issues, challenges, techniques, and future directions," *Journal of Imaging*, vol. 5, no. 4, p. 48, 2019.

[26] B. Grabowski, W. Masarczyk, P. Głomb, and A. Mendys, "Automatic pigment identification from hyperspectral data," *Journal of Cultural Heritage*, vol. 31, pp. 1–12, 2018.

[27] B. Brunetti, C. Miliani, F. Rosi, B. Doherty, L. Monico, A. Romani, and A. Sgamellotti, "Non-invasive investigations of paintings by portable instrumentation: the MOLAB experience," *Analytical Chemistry for Cultural Heritage*, pp. 41–75, 2017.

[28] M. P. Colombini, I. Degano, and A. Nevin, "Analytical chemistry for the study of paintings and the detection of forgeries," *Foundations*, 2016.

[29] B. Borg, M. Dunn, A. Ang, and C. Villis, "The application of state-of-the-art technologies to support artwork conservation: Literature review," *Journal of Cultural Heritage*, vol. 44, pp. 239–259, 2020.

[30] G. Nehring, O. Bonnerot, M. Gerhardt, M. Krutzsch, and I. Rabin, "Looking for the missing link in the evolution of black inks," *Archaeological and Anthropological Sciences*, vol. 13, pp. 1–10, 2021.

[31] R. Córdoba de la Llave, "Interdisciplinary exploration of medieval technical manuscripts from the iberian peninsula," *Journal of Medieval Iberian Studies*, vol. 14, no. 1, pp. 96–108, 2022.

[32] A. Potthast, U. Henniges, and G. Banik, "Iron gall ink-induced corrosion of cellulose: aging, degradation and stabilization. part 1: model paper studies," *Cellulose*, vol. 15, pp. 849–859, 2008.

[33] C. Crespo and V. Viñas, *La preservación y restauración de documentos y libros en papel: un estudio del RAMP con directrices*. UNESCO, 1984.

[34] C. A. Mitchell, "INKS. Lecture I," *Journal of the Royal Society of Arts*, vol. 70, no. 3637, pp. 647–660, 1922.

[35] I. Rabin, "Material studies of historic inks: transition from carbon to iron-gall inks," in *Traces of ink*, pp. 70–78, Brill, 2021.

[36] C. Colini, O. Hahn, O. Bonnerot, S. Steger, Z. Cohen, T. Ghigo, T. Christiansen, M. Bicchieri, P. Biocca, M. Krutzsch, *et al.*, "The quest for the mixed inks," *manuscript cultures*, vol. 2018, no. 11, pp. 41–48, 2018.

[37] M. Schreiner, B. Frühmann, D. Jembrih-Simbürger, and R. Linke, "X-rays in art and archaeology: An overview," *Powder Diffraction*, vol. 19, no. 1, pp. 3–11, 2004.

[38] S. Bergeon, "Painting technique: priming, coloured paint film and varnish," *Scientific Examination of Easel Paintings*, pp. 35–62, 1986.

[39] G. Bitossi, R. Giorgi, M. Mauro, B. Salvadori, and L. Dei, "Spectroscopic techniques in cultural heritage conservation: a survey," *Applied Spectroscopy Reviews*, vol. 40, no. 3, pp. 187–228, 2005.

[40] I. Degano, E. Ribechini, F. Modugno, and M. P. Colombini, "Analytical methods for the characterization of organic dyes in artworks and in historical textiles," *Applied Spectroscopy Reviews*, vol. 44, no. 5, pp. 363–410, 2009.

[41] M. Sawczak, A. Kamińska, G. Rabczuk, M. Ferretti, R. Jendrzejewski, and G. Śliwiński, "Complementary use of the raman and xrf techniques for non-destructive analysis of historical paint layers," *Applied surface science*, vol. 255, no. 10, pp. 5542–5545, 2009.

[42] F. Rosi, A. Burnstock, K. J. Van den Berg, C. Miliani, B. G. Brunetti, and A. Sgamellotti, "A non-invasive XRF study supported by multivariate statistical analysis and reflectance FTIR to assess the composition of modern painting materials," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 71, no. 5, pp. 1655–1662, 2009.

[43] A. L. de Queiroz Baddini, J. L. V. de Paula Santos, R. R. Tavares, L. S. de Paula, H. da Costa Araújo Filho, and R. P. Freitas, "PLS-DA and data fusion of visible Reflectance, XRF and FTIR spectroscopy in the classification of mixed historical pigments," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 265, p. 120384, 2022.

[44] B. Hochleitner, V. Desnica, M. Mantler, and M. Schreiner, "Historical pigments: a collection analyzed with x-ray diffraction analysis and x-ray fluorescence analysis in order to create a database," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 58, no. 4, pp. 641–649, 2003.

[45] M. Eveno, B. Moignard, and J. Castaing, "Portable apparatus for in situ x-ray diffraction and fluorescence analyses of artworks," *Microscopy and Microanalysis*, vol. 17, no. 5, pp. 667–673, 2011.

[46] L. B. Brostoff, S. Centeno, P. Ropret, P. Bythrow, and F. Pottier, "Combined x-ray diffraction and raman identification of synthetic organic pigments in works of art: From powder samples to artists' paints," *Analytical Chemistry*, vol. 81, no. 15, pp. 6096–6106, 2009.

[47] D. Buti, F. Rosi, B. G. Brunetti, and C. Miliani, "In-situ identification of copper-based green pigments on paintings and manuscripts by reflection FTIR," *Analytical and bioanalytical chemistry*, vol. 405, pp. 2699–2711, 2013.

[48] N. C. Scherrer, Z. Stefan, D. Francoise, F. Annette, and K. Renate, "Synthetic organic pigments of the 20th and 21st century relevant to artist's paints: Raman spectra reference collection," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 73, no. 3, pp. 505–524, 2009.

[49] P. Sarrazin, G. Chiari, and M. Gailhanou, "A portable non-invasive XRD-XRF instrument for the study of art objects," *Advances in X-ray analysis*, vol. 52, pp. 175–86, 2008.

[50] P. Ricciardi, S. Legrand, G. Bertolotti, and K. Janssens, "Macro X-ray fluorescence (MA-XRF) scanning of illuminated manuscript fragments: potentialities and challenges," *Microchemical Journal*, vol. 124, pp. 785–791, 2016.

[51] P. Vandenabeele and A. Rousaki, "Developing macro-Raman mapping as a tool for studying the pigment distribution of art objects," *Analytical Chemistry*, vol. 93, no. 46, pp. 15390–15400, 2021.

[52] S. Legrand, P. Ricciardi, L. Nodari, and K. Janssens, "Non-invasive analysis of a 15th century illuminated manuscript fragment: point-based vs imaging spectroscopy," *Microchemical Journal*, vol. 138, pp. 162–172, 2018.

[53] European Committee for Standardization (CEN), "Solar energy—vocabulary," standard, European Committee for Standardization (CEN), Brussels, Oct. 1999.

[54] F. I. G. Rawlins, "Studies in the colorimetry of paintings," *Technical studies in the field of the fine arts*, vol. 4, no. 4, pp. 179–186, 1936.

[55] N. F. Barnes, "A spectrophotometric study of artists' pigments," *Technical Studies in the Field of the Fine Arts*, vol. 7, no. 3, 1939.

[56] M. Picollo, M. Aceto, and T. Vitorino, "UV-Vis spectroscopy," *Physical sciences reviews*, vol. 4, no. 4, p. 20180008, 2019.

[57] M. Bacci, S. Baronti, A. Casini, P. Castagna, R. Linari, A. Orlando, M. Picollo, and B. Radicati, "Detection of alteration products in artworks by non-destructive spectroscopic analysis," *MRS Online Proceedings Library (OPL)*, vol. 352, p. 153, 1995.

[58] M. Bacci, A. Casini, C. Cucci, M. Picollo, B. Radicati, and M. Vervat, "Non-invasive spectroscopic measurements on the Il ritratto della figliastra by Giovanni Fattori: identification of pigments and colourimetric analysis," *Journal of cultural heritage*, vol. 4, no. 4, pp. 329–336, 2003.

[59] M. Leona and J. Winter, "Fiber optics reflectance spectroscopy: a unique tool for the investigation of Japanese paintings," *Studies in Conservation*, vol. 46, no. 3, pp. 153–162, 2001.

[60] M. Picollo, M. Bacci, A. Casini, F. Lotti, S. Porcinai, B. Radicati, and L. Stefani, "Fiber optics reflectance spectroscopy: a non-destructive technique for the analysis of works of art," *Optical sensors and microsystems: new concepts, materials, technologies*, pp. 259–265, 2000.

[61] A. M. C. A. N. 75, "UV-visible-NIR reflectance spectrophotometry in cultural heritage: Background paper," *Analytical Methods*, vol. 8, no. 30, pp. 5894–5896, 2016.

[62] A. Pallipurath, J. Skelton, P. Ricciardi, S. Bucklow, and S. Elliott, "Multivariate analysis of combined raman and fibre-optic reflectance spectra for the identification of binder materials in simulated medieval paints," *Journal of Raman spectroscopy*, vol. 44, no. 6, pp. 866–874, 2013.

[63] E. Ciliberto and G. Spoto, *Modern analytical methods in art and archaeology*. Wiley, 2000.

[64] J. Striova, A. Dal Fovo, and R. Fontana, "Reflectance imaging spectroscopy in heritage science," *La Rivista del Nuovo Cimento*, vol. 43, pp. 515–566, 2020.

[65] U. Siripatrawan and Y. Makino, "Hyperspectral imaging coupled with machine learning for classification of anthracnose infection on mango fruit," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 309, p. 123825, 2024.

[66] A. Bhargava, A. Sachdeva, K. Sharma, M. H. Alsharif, P. Uthansakul, and M. Uthansakul, "Hyperspectral imaging and its applications: A review," *Heliyon*, vol. 10, no. 12, 2024.

[67] Y. W. Wang, N. P. Reder, S. Kang, A. K. Glaser, and J. T. Liu, "Multiplexed optical imaging of tumor-directed nanoparticles: a review of imaging systems and approaches," *Nanotheranostics*, vol. 1, no. 4, p. 369, 2017.

[68] Y. Wu, Z. Xu, W. Yang, Z. Ning, and H. Dong, "Review on the application of hyperspectral imaging technology of the exposed cortex in cerebral surgery," *Frontiers in Bioengineering and Biotechnology*, vol. 10, p. 906728, 2022.

[69] F. G. France, "Spectral imaging for preservation documentation," in *Archiving Conference*, vol. 13, pp. 2–5, Society for Imaging Science and Technology, 2016.

[70] J. Dyer, G. Verri, and J. Cupitt, *Multispectral imaging in reflectance and photo-induced luminscence modes: a user manual*. British Museum London, UK, 2013.

[71] H. Liang, "Advances in multispectral and hyperspectral imaging for archaeology and art conservation," *Applied Physics A*, vol. 106, pp. 309–323, 2012.

[72] C. Cucci, J. K. Delaney, and M. Picollo, "Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts," *Accounts of chemical research*, vol. 49, no. 10, pp. 2070–2079, 2016.

[73] L. G. Hamey, "Radiometric camera calibration," *Wiley Encyclopedia of Computer Science and Engineering*, pp. 1–10, 2007.

[74] A. Burmester, J. Cupitt, H. Derrien, N. Dessipris, A. Hamber, K. Martinez, M. Muller, and D. Saunders, "The examination of paintings by digital image analysis," in *3rd International Conference on Non Destructive Testing, Microanalytical Methods and Environmental Evaluation for Study and Conservation of Works of Art Rome*, pp. 199–214, The International Society for Optical Engineering, 1992.

[75] S. Baronti, A. Casini, F. Lotti, and S. Porcinai, "Principal component analysis of visible and near-infrared multispectral images of works of art," *Chemometrics and Intelligent Laboratory Systems*, vol. 39, no. 1, pp. 103–114, 1997.

[76] A. Casini, F. Lotti, M. Picollo, L. Stefani, and E. Buzzegoli, "Image spectroscopy mapping technique for noninvasive analysis of paintings," *Studies in conservation*, vol. 44, no. 1, pp. 39–48, 1999.

[77] C. Balas, V. Papadakis, N. Papadakis, A. Papadakis, E. Vazgiouraki, and G. Themelis, "A novel hyper-spectral imaging apparatus for the non-destructive analysis of objects of artistic and historic value," *Journal of Cultural Heritage*, vol. 4, pp. 330–337, 2003.

[78] A. Casini, M. Bacci, C. Cucci, F. Lotti, S. Porcinai, M. Picollo, B. Radicati, M. Poggesi, and L. Stefani, "Fiber optic reflectance spectroscopy and hyper-spectral image spectroscopy: two integrated techniques for the study of the Madonna dei Fusi," in *Optical methods for arts and archaeology*, vol. 5857, pp. 177–184, SPIE, 2005.

[79] P. Cotte and D. Dupraz, "Spectral imaging of leonardo da vinci's mona lisa: An authentic smile at 1523 dpi with additional infrared data," in *Archiving Conference*, vol. 3, pp. 228–235, Society of Imaging Science and Technology, 2006.

[80] F. H. Imai, M. R. Rosen, and R. S. Berns, "Multi-spectral Imaging of a van Gogh's Self-portrait at the National Gallery of Art Washington DC.," in *PICS*, pp. 185–189, 2001.

[81] J. K. Delaney, K. A. Dooley, A. Van Loon, and A. Vandivere, "Mapping the pigment distribution of Vermeer's Girl with a Pearl Earring," *Heritage Science*, vol. 8, pp. 1–16, 2020.

[82] F. G. France, "Advanced spectral imaging for noninvasive microanalysis of cultural heritage materials: review of application to documents in the US Library of Congress," *Applied spectroscopy*, vol. 65, no. 6, pp. 565–574, 2011.

[83] J. K. Delaney, J. G. Zeibel, M. Thoury, R. Littleton, M. Palmer, K. M. Morales, E. R. de La Rie, and A. Hoenigswald, "Visible and infrared imaging spectroscopy of Picasso's Harlequin musician: mapping and identification of artist materials in situ," *Applied spectroscopy*, vol. 64, no. 6, pp. 584–594, 2010.

[84] C. Balas, G. Epitropou, A. Tsapras, and N. Hadjinicolaou, "Hyperspectral imaging and spectral classification for pigment identification and mapping in paintings by El Greco and his workshop," *Multimedia Tools and Applications*, vol. 77, pp. 9737–9751, 2018.

[85] C.-a. Wei, J. Li, and S. Liu, "Applications of visible spectral imaging technology for pigment identification of colored relics," *Heritage Science*, vol. 12, no. 1, p. 321, 2024.

[86] M. Picollo, C. Cucci, A. Casini, and L. Stefani, "Hyper-spectral imaging technique in the cultural heritage field: New possible scenarios," *Sensors*, vol. 20, no. 10, p. 2843, 2020.

[87] G. Trumpy and S. Chatterjee, "A spectral approach to digitally restore a faded agfacolor print from 1945," in *Archiving Conference*, vol. 19, pp. 101–105, Society for Imaging Science and Technology, 2022.

[88] W. A. Christens-Barry, K. Boydston, F. G. France, K. T. Knox, R. L. Easton Jr, and M. B. Toth, "Camera system for multispectral imaging of documents," in *Sensors, Cameras, and Systems for Industrial/Scientific Applications X*, vol. 7249, pp. 32–41, SPIE, 2009.

[89] C. Cucci, O. De Pascale, and G. S. Senesi, "Assessing laser cleaning of a limestone monument by fiber optics reflectance spectroscopy (FORS) and visible and near-Infrared (VNIR) hyperspectral imaging (HSI)," *Minerals*, vol. 10, no. 12, p. 1052, 2020.

[90] E. Matoušková, K. Kovářová, M. Cihla, and J. Hodač, "Monitoring biological degradation of historical stone using hyperspectral imaging," *European Journal of Remote Sensing*, vol. 57, no. 1, p. 2220565, 2024.

[91] M. Martínez Domingo, E. Valero, R. Huertas, M. Durbán, T. Espejo, and R. Blanc, "Spectral information to get beyond color in the analysis of water-soluble varnish degradation," *Heritage Science*, vol. 7, pp. 1–16, 2019.

[92] CHSOS, "Technical photography (TP)." `https://chsopensource.org/1-technical-photography-tp/`, 2025. Accessed: February 18, 2025.

[93] P. Shor, M. Manfredi, G. H. Bearman, E. Marengo, K. Boydston, and W. A. Christens-Barry, "The leon levy dead sea scrolls digital library: The digitization project of the dead sea scrolls," *Journal of Eastern Mediterranean Archaeology & Heritage Studies*, vol. 2, no. 2, pp. 71–89, 2014.

[94] C. Bassani, R. M. Cavalli, R. Goffredo, A. Palombo, S. Pascucci, and S. Pignatti, "Specific spectral bands for different land cover contexts to improve the efficiency of remote sensing archaeological prospection: The Arpi case study," *Journal of Cultural Heritage*, vol. 10, pp. e41–e48, 2009.

[95] L. P. Pardo, P. Dryburgh, E. Biggs, M. Vermeulen, P. Crooks, A. Gibson, M. Fort, C. Vlachou-Mogire, M. Bertasa, J. R. Gilchrist, *et al.*, "Advanced imaging to recover illegible text in historic documents. the challenge of past chemical treatments for ink enhancement," *Journal of Cultural Heritage*, vol. 68, pp. 342–353, 2024.

[96] E. Kirchner, I. van der Lans, F. Ligterink, E. Hendriks, and J. Delaney, "Digitally reconstructing Van Gogh's field with irises near Arles. Part 1: varnish," *Color Research & Application*, vol. 43, no. 2, pp. 150–157, 2018.

[97] Z. Zeng, S. Qiu, P. Zhang, X. Tang, S. Li, X. Liu, and B. Hu, "Virtual restoration of ancient tomb murals based on hyperspectral imaging," *Heritage Science*, vol. 12, no. 1, p. 410, 2024.

[98] K. Imai, J. K. Delaney, S. Webster-Cook, and R. Jiménez Garnica, "Pablo Picasso's Mother and Child by the Sea (1902): Report on the hyperspectral near-infrared reflectance imaging survey of Picasso's newspaper use," *SN Applied Sciences*, vol. 2, no. 9, p. 1474, 2020.

[99] G. Trumpy, J. Y. Hardeberg, S. George, and B. Flueckiger, "Multispectral capture of film colors with LED," in *Colour Photography and Film: Sharing knowledge of analysis, preservation, conservation, migration of analogue and digital materials*, 2021.

[100] C. Cucci, A. Casini, L. Stefani, B. Cattaneo, and M. Picollo, "A novel transmittance Vis–NIR hyper-spectral imaging scanner for analysis of photographic negatives: A potential tool for photography conservation," *Sensors*, vol. 23, no. 7, p. 3562, 2023.

[101] M. Á. Martínez-Domingo, A. I. Calero Castillo, E. Vivar García, and E. M. Valero, "Evaluation of cleaning processes using colorimetric and spectral data for the removal of layers of limewash from medieval plasterwork," *Sensors*, vol. 20, no. 24, p. 7147, 2020.

[102] E. M. Valero, E. Vivar, M. Á. Martínez-Domingo, and A. I. Calero, "Unmixing-based cleaning methods evaluation for re-polychromated plasterwork," *Journal of Cultural Heritage*, vol. 66, pp. 10–19, 2024.

[103] M. Á. Martínez, A. I. Calero, and E. M. Valero, "Colorimetric and spectral data analysis of consolidants used for preservation of medieval plasterwork," *Journal of Cultural Heritage*, vol. 42, pp. 64–71, 2020.

[104] W. Kautek, S. Pentzien, D. Mueller-Hess, K. Troschke, and R. Teule, "Probing the limits of paper and parchment laser cleaning by multispectral imaging," in *Laser Techniques and Systems in Art Conservation*, vol. 4402, pp. 130–138, SPIE, 2001.

[105] R. S. Berns, J. Krueger, and M. Swicklik, "Multiple pigment selection for inpainting using visible reflectance spectrophotometry," *Studies in conservation*, vol. 47, no. 1, pp. 46–61, 2002.

[106] L. Pronti, M. Romani, G. Verona-Rinati, O. Tarquini, F. Colao, M. Colapietro, A. Pifferi, M. Cestelli-Guidi, and M. Marinelli, "Post-processing of VIS, NIR, and SWIR multispectral images of paintings. New discovery on the the drunkenness of Noah, painted by Andrea Sacchi, stored at Palazzo Chigi (Ariccia, Rome)," *Heritage*, vol. 2, no. 3, pp. 2275–2286, 2019.

[107] R. L. Easton, W. A. Christens-Barry, and K. T. Knox, "Spectral image processing and analysis of the Archimedes Palimpsest," in *2011 19th European Signal Processing Conference*, pp. 1440–1444, IEEE, 2011.

[108] C. Cucci, E. K. Webb, A. Casini, M. Ginanni, E. Prandi, L. Stefani, T. Vitorino, and M. Picollo, "Short-wave infrared reflectance hyperspectral imaging for painting investigations: A methodological study," *Journal of the American Institute for Conservation*, vol. 58, no. 1-2, pp. 16–36, 2019.

[109] K. Rapantzikos and C. Balas, "Hyperspectral imaging: potential in non-destructive analysis of palimpsests," in *IEEE International Conference on Image Processing 2005*, vol. 2, pp. II–618, IEEE, 2005.

[110] D. Saunders and J. Cupitt, "Image processing at the National Gallery: The VASARI project," *National Gallery technical bulletin*, vol. 14, no. 1, pp. 72–85, 1993.

[111] M. Á. Martínez-Domingo, M. Melgosa, K. Okajima, V. J. Medina, and F. J. Collado-Montero, "Spectral image processing for museum lighting using CIE LED illuminants," *Sensors*, vol. 19, no. 24, p. 5400, 2019.

[112] M. Á. Martínez-Domingo, A. B. López-Baldomero, M. Tejada-Casado, M. Melgosa, and F. J. Collado-Montero, "Colorimetric evaluation of a reintegration via spectral imaging—case study: Nasrid tiling panel from the Alhambra of Granada (Spain)," *Sensors*, vol. 24, no. 12, p. 3872, 2024.

[113] G. Trumpy, D. Conover, L. Simonot, M. Thoury, M. Picollo, and J. K. Delaney, "Experimental study on merits of virtual cleaning of paintings with aged varnish," *Optics express*, vol. 23, no. 26, pp. 33836–33848, 2015.

[114] Z. Wang, D. Lu, D. Zhang, M. Sun, and Y. Zhou, "Fake modern Chinese painting identification based on spectral–spatial feature fusion on hyperspectral image," *Multidimensional Systems and Signal Processing*, vol. 27, pp. 1031–1044, 2016.

[115] S. J. Kim, F. Deng, and M. S. Brown, "Visual enhancement of old documents with hyperspectral imaging," *Pattern Recognition*, vol. 44, no. 7, pp. 1461–1469, 2011.

[116] M. Lettner, M. Diem, R. Sablatnig, and H. Miklas, "Registration and enhancing of multispectral manuscript images," in *2008 16th European Signal Processing Conference*, pp. 1–5, IEEE, 2008.

[117] G. A. Ware, D. M. Chabries, R. W. Christiansen, and C. E. Martin, "Multispectral document enhancement: Ancient carbonized scrolls," in *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the*

Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120), vol. 6, pp. 2486–2488, IEEE, 2000.

[118] R. M. Cavalli, F. Colosi, A. Palombo, S. Pignatti, and M. Poscolieri, "Remote hyperspectral imagery as a support to archaeological prospection," *Journal of Cultural Heritage*, vol. 8, no. 3, pp. 272–283, 2007.

[119] S. Al Maadeed, S. Kunhoth, and A. Bouridane, "Digitization and spectral analysis of historical manuscripts," in *2014 World Symposium on Computer Applications & Research (WSCAR)*, pp. 1–6, IEEE, 2014.

[120] S. Tanner and G. Bearman, "Digitizing the dead sea scrolls," in *Archiving Conference*, vol. 6, pp. 119–123, Society of Imaging Science and Technology, 2009.

[121] R. L. Easton Jr, K. T. Knox, W. A. Christens-Barry, K. Boydston, M. B. Toth, D. Emery, and W. Noel, "Standardized system for multispectral imaging of palimpsests," in *Computer Vision and Image Analysis of Art*, vol. 7531, pp. 106–116, SPIE, 2010.

[122] M. Attas, E. Cloutis, C. Collins, D. Goltz, C. Majzels, J. R. Mansfield, and H. H. Mantsch, "Near-infrared spectroscopic imaging in art conservation: investigation of drawing constituents," *Journal of Cultural Heritage*, vol. 4, no. 2, pp. 127–136, 2003.

[123] A. Cosentino, "Fors spectral database of historical pigments in different binders," *E Conserv. J*, vol. 2, pp. 54–65, 2014.

[124] H. Deborah, "Hyperspectral pigment dataset," in *2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–5, IEEE, 2022.

[125] S. Vishnu, R. R. Nidamanuri, and R. Bremananth, "Spectral material mapping using hyperspectral imagery: a review of spectral matching and library search methods," *Geocarto international*, vol. 28, no. 2, pp. 171–190, 2013.

[126] D. J. Mandal, M. Pedersen, S. George, H. Deborah, and C. Boust, "An experiment-based comparative analysis of pigment classification algorithms using hyperspectral imaging," *Journal of Imaging Science and Technology*, vol. 67, no. 3, pp. 030403–1–030403–18, 2023.

[127] J. K. Delaney, P. Ricciardi, L. D. Glinsman, M. Facini, M. Thoury, M. Palmer, and E. R. d. l. Rie, "Use of imaging spectroscopy, fiber optic reflectance spectroscopy, and X-ray fluorescence to map and identify pigments in illuminated manuscripts," *Studies in Conservation*, vol. 59, no. 2, pp. 91–101, 2014.

[128] F. Gabrieli, K. Dooley, M. Facini, and J. Delaney, "Near-UV to mid-IR reflectance imaging spectroscopy of paintings on the macroscale," *Science advances*, vol. 5, no. 8, p. eaaw7794, 2019.

[129] B. M. Devassy, S. George, and J. Y. Hardeberg, "Comparison of ink classification capabilities of classic hyperspectral similarity features," in *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, vol. 8, pp. 25–30, IEEE, 2019.

[130] D. J. Mandal, S. George, M. Pedersen, and C. Boust, "Influence of acquisition parameters on pigment classification using hyperspectral imaging," in *Color and Imaging Conference*, vol. 29, pp. 334–346, Society for Imaging Science and Technology, 2021.

[131] S. George and J. Y. Hardeberg, "Ink classification and visualisation of historical manuscripts: Application of hyperspectral imaging," in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1131–1135, IEEE, 2015.

[132] A. Cesaratto, A. Nevin, G. Valentini, L. Brambilla, C. Castiglioni, L. Toniolo, M. Fratelli, and D. Comelli, "A novel classification method for multispectral imaging combined with portable Raman spectroscopy for the analysis of a painting by Vincent Van Gogh," *Applied spectroscopy*, vol. 67, no. 11, pp. 1234–1241, 2013.

[133] D. Comelli, A. Nevin, G. Valentini, I. Osticioli, E. M. Castellucci, L. Toniolo, D. Gulotta, and R. Cubeddu, "Insights into masolino's wall paintings in castiglione olona: Advanced reflectance and fluorescence imaging analysis," *Journal of cultural heritage*, vol. 12, no. 1, pp. 11–18, 2011.

[134] T. Kleynhans, D. W. Messinger, and J. K. Delaney, "Towards automatic classification of diffuse reflectance image cubes from paintings collected with hyperspectral cameras," *Microchemical Journal*, vol. 157, p. 104934, 2020.

[135] M. Romani, G. Capobianco, L. Pronti, F. Colao, C. Seccaroni, A. Puiu, A. Felici, G. Verona-Rinati, M. Cestelli-Guidi, A. Tognacci, *et al.*, "Analytical chemistry approach in cultural heritage: the case of Vincenzo Pasqualoni's wall paintings in S. Nicola in Carcere (Rome)," *Microchemical Journal*, vol. 156, p. 104920, 2020.

[136] J. Cai, H. Chatoux, C. Boust, and A. Mansouri, "Comparison of linear unmixing methods on paintings data set," in *ORASIS 2021*, 2021.

[137] E. Pouyet, T. Miteva, N. Rohani, and L. de Viguerie, "Artificial intelligence for pigment classification task in the short-wave infrared range," *Sensors*, vol. 21, no. 18, p. 6150, 2021.

[138] R. Radpour, G. A. Gates, I. Kakoulli, and J. K. Delaney, "Identification and mapping of ancient pigments in a Roman Egyptian funerary portrait by application of reflectance and luminescence imaging spectroscopy," *Heritage Science*, vol. 10, no. 1, p. 8, 2022.

[139] C. Vlachou-Mogire, J. Danskin, J. R. Gilchrist, and K. Hallett, "Mapping materials and dyes on historic tapestries using hyperspectral imaging," *Heritage*, vol. 6, no. 3, pp. 3159–3182, 2023.

[140] H. Deborah, N. Richard, and J. Y. Hardeberg, "A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3224–3234, 2015.

[141] J. K. Delaney, P. Ricciardi, L. Glinsman, M. Palmer, and J. Burke, "Use of near infrared reflectance imaging spectroscopy to map wool and silk fibres in historic tapestries," *Analytical Methods*, vol. 8, no. 44, pp. 7886–7890, 2016.

[142] H. Deborah, S. George, and J. Y. Hardeberg, "Pigment mapping of the scream (1893) based on hyperspectral imaging," in *Image and Signal Processing: 6th International Conference, ICISP 2014, Cherbourg, France, June 30–July 2, 2014. Proceedings 6*, pp. 247–256, Springer, 2014.

[143] P. Ricciardi, J. K. Delaney, M. Facini, and L. Glinsman, "Use of imaging spectroscopy and in situ analytical methods for the characterization of the materials and techniques of 15th century illuminated manuscripts," *Journal of the American Institute for Conservation*, vol. 52, no. 1, pp. 13–29, 2013.

[144] A. Dal Fovo, M. Morello, A. Mazzinghi, C. Toso, M. Galeotti, and R. Fontana, "Spectral mapping techniques for the stratigraphic and compositional characterisation of a 16th-century painting," *Heritage*, vol. 7, no. 3, pp. 1320–1333, 2024.

[145] Z. Mingming, M. Lei, and L. Feimo, "Representative spectral extraction approach for hyperspectral images based on mixed spectral similarity metric," in *2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, pp. 1675–1680, IEEE, 2019.

[146] J. Ding, X. Li, and L. Huang, "A novel method for spectral similarity measure by fusing shape and amplitude features," *Journal of Engineering Science & Technology Review*, vol. 8, no. 5, 2015.

[147] A. Chen, R. Jesus, and M. Vilarigues, "Identification of pure painting pigment using machine learning algorithms," in *Artificial Intelligence in Music, Sound, Art and Design: 10th International Conference, EvoMUSART 2021, Held as Part of*

EvoStar 2021, Virtual Event, April 7–9, 2021, Proceedings 10, pp. 52–64, Springer, 2021.

[148] M. Kubik, "Hyperspectral imaging: a new technique for the non-invasive study of artworks," in *Physical techniques in the study of art, archaeology and cultural heritage*, vol. 2, pp. 199–259, Elsevier, 2007.

[149] H. Deborah, N. Richard, and J. Y. Hardeberg, "On the quality evaluation of spectral image processing algorithms," in *2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems*, pp. 133–140, IEEE, 2014.

[150] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: an overview," *arXiv preprint arXiv:2008.05756*, 2020.

[151] L. Liu, T. Miteva, G. Delnevo, S. Mirri, P. Walter, L. de Viguerie, and E. Pouyet, "Neural networks for hyperspectral imaging of historical paintings: a practical review," *Sensors*, vol. 23, no. 5, p. 2419, 2023.

[152] F. Liggins, A. Vichi, W. Liu, A. Hogg, S. Kogou, J. Chen, and H. Liang, "Hyperspectral imaging solutions for the non-invasive detection and automated mapping of copper trihydroxychlorides in ancient bronze," *Heritage Science*, vol. 10, no. 1, p. 142, 2022.

[153] L. Butler, S. Kogou, A. Hogg, Y. Li, A. Vichi, A. Gallop, and H. Liang, "Machine learning analysis of complementary multimodal spectral imaging data from a large manuscript collection," in *Optics for Arts, Architecture, and Archaeology VIII*, vol. 11784, p. 1178405, SPIE, 2021.

[154] Z. Jianxin, Z. Kangping, W. Junkai, and H. Xudong, "Color segmentation and extraction of yarn-dyed fabric based on a hyperspectral imaging system," *Textile Research Journal*, vol. 91, no. 7-8, pp. 729–742, 2021.

[155] N. Rohani, J. Salvant, S. Bahaadini, O. Cossairt, M. Walton, and A. Katsaggelos, "Automatic pigment identification on roman egyptian paintings by using sparse modeling of hyperspectral images," in *2016 24th European signal processing conference (EUSIPCO)*, pp. 2111–2115, IEEE, 2016.

[156] D. Bai, D. W. Messinger, and D. Howell, "A hyperspectral imaging spectral unmixing and classification approach to pigment mapping in the Gough & Selden Maps," *Journal of the American Institute for Conservation*, vol. 58, no. 1-2, pp. 69–89, 2019.

[157] G. Capobianco, L. Pronti, E. Gorga, M. Romani, M. Cestelli-Guidi, S. Serranti, and G. Bonifazi, "Methodological approach for the automatic discrimination of pictorial materials using fused hyperspectral imaging data from the visible to mid-infrared range coupled with machine learning methods," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 304, p. 123412, 2024.

[158] A. Chen, R. Jesus, and M. Vilarigues, "Convolutional neural network-based pure paint pigment identification using hyperspectral images," in *Proceedings of the 3rd ACM International Conference on Multimedia in Asia*, pp. 1–7, 2021.

[159] J. Huang, H. He, R. Lv, G. Zhang, Z. Zhou, and X. Wang, "Non-destructive detection and classification of textile fibres based on hyperspectral imaging and 1D-CNN," *Analytica Chimica Acta*, vol. 1224, p. 340238, 2022.

[160] J. Blažek, J. Striová, R. Fontana, and B. Zitová, "Improvement of the visibility of concealed features in artwork NIR reflectograms by information separation," *Digital Signal Processing*, vol. 60, pp. 140–151, 2017.

[161] N. Rohani, E. Pouyet, M. Walton, O. Cossairt, and A. K. Katsaggelos, "Nonlinear unmixing of hyperspectral datasets for the study of painted works of art," *Angewandte Chemie*, vol. 130, no. 34, pp. 11076–11080, 2018.

[162] T. Kleynhans, C. M. Schmidt Patterson, K. A. Dooley, D. W. Messinger, and J. K. Delaney, "An alternative approach to mapping pigments in paintings with hyperspectral reflectance image cubes using artificial intelligence," *Heritage Science*, vol. 8, pp. 1–16, 2020.

[163] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[164] I. M. Ciortan, T. G. Poulsson, S. George, and J. Y. Hardeberg, "Tensor decomposition for painting analysis. part 1: pigment characterization," *Heritage Science*, vol. 11, no. 1, p. 76, 2023.

[165] J. Van der Weerd, A. Van Loon, and J. J. Boon, "FTIR studies of the effects of pigments on the aging of oil," *Studies in conservation*, vol. 50, no. 1, pp. 3–22, 2005.

[166] D. Saunders and J. Kirby, "The effect of relative humidity on artists' pigments," *National Gallery Technical Bulletin*, vol. 25, pp. 62–72, 2004.

[167] L. Simonot and M. Elias, "Color change due to surface state modification," *Color Research & Application*, vol. 28, no. 1, pp. 45–49, 2003.

[168] H. Deborah, M. O. Ulfarsson, and J. Sigurdsson, "Fully constrained least squares linear spectral unmixing of The Scream (Verso, 1893)," in *2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–5, IEEE, 2021.

[169] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Imaging spectrometry V*, vol. 3753, pp. 266–275, SPIE, 1999.

[170] Y. Li and A. Ngom, "The non-negative matrix factorization toolbox for biological data mining," *Source code for biology and medicine*, vol. 8, pp. 1–15, 2013.

[171] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Deep generative endmember modeling: An application to unsupervised spectral unmixing," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 374–384, 2019.

[172] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *2009 First workshop on hyperspectral image and signal processing: Evolution in remote sensing*, pp. 1–4, IEEE, 2009.

[173] X. Song, X. Jiang, and X. Rui, "Spectral unmixing using linear unmixing under spatial autocorrelation constraints," in *2010 IEEE International Geoscience and Remote Sensing Symposium*, pp. 975–978, IEEE, 2010.

[174] S. Stagakis, T. Vanikiotis, and O. Sykioti, "Estimating forest species abundance through linear unmixing of CHRIS/PROBA imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 119, pp. 79–89, 2016.

[175] M. Nie, Z. Liu, X. Li, Q. Wu, B. Tang, X. Xiao, Y. Sun, J. Chang, and C. Zheng, "Novel method for hyperspectral unmixing: fuzzy c-means unmixing," *Sensor Review*, vol. 36, no. 2, pp. 184–192, 2016.

[176] G. Zhang, S. Mei, B. Xie, M. Ma, Y. Zhang, Y. Feng, and Q. Du, "Spectral variability augmented sparse unmixing of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[177] F. Grillini, J.-B. Thomas, and S. George, "Comparison of imaging models for spectral unmixing in oil painting," *Sensors*, vol. 21, no. 7, p. 2471, 2021.

[178] F. Grillini, J.-B. Thomas, and S. George, "Linear, subtractive and logarithmic optical mixing models in oil painting.," in *CVCS*, 2020.

[179] S. Lyu, D. Meng, M. Hou, S. Tian, C. Huang, and J. Mao, "Nonlinear mixing characteristics of reflectance spectra of typical mineral pigments," *Minerals*, vol. 11, no. 6, p. 626, 2021.

[180] D. Duncan, "The colour of pigment mixtures," *Proceedings of the Physical Society*, vol. 52, no. 3, p. 390, 1940.

[181] H. Yang, S. Zhu, and N. Pan, "On the Kubelka—Munk single-constant/two-constant theories," *Textile Research Journal*, vol. 80, no. 3, pp. 263–270, 2010.

[182] S. Kogou, A. Lucian, S. Bellesia, L. Burgio, K. Bailey, C. Brooks, and H. Liang, "A holistic multimodal approach to the non-invasive analysis of watercolour paintings," *Applied Physics A*, vol. 121, pp. 999–1014, 2015.

[183] N. Rohani, E. Pouyet, M. Walton, O. Cossairt, and A. K. Katsaggelos, "Pigment unmixing of hyperspectral images of paintings using deep neural networks," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3217–3221, IEEE, 2019.

[184] L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3890–3905, 2016.

[185] J. Nascimento and G. Martín, "Nonlinear spectral unmixing," in *Data Handling in Science and Technology*, vol. 32, pp. 151–166, Elsevier, 2019.

[186] K. Fukumoto, N. Tsumura, and R. Berns, "Estimating pigment concentrations from spectral images using an encoder decoder neural network.," *Journal of Imaging Science & Technology*, vol. 64, no. 3, 2020.

[187] L. Qi, J. Li, Y. Wang, M. Lei, and X. Gao, "Deep spectral convolution network for hyperspectral image unmixing with spectral library," *Signal Processing*, vol. 176, p. 107672, 2020.

[188] L. Qi, F. Gao, J. Dong, X. Gao, and Q. Du, "SSCU-Net: Spatial–spectral collaborative unmixing network for hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[189] B. Rasti, B. Koirala, P. Scheunders, and J. Chanussot, "Misicnet: Minimum simplex convolutional network for deep hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[190] J. K. Delaney, M. Thoury, J. G. Zeibel, P. Ricciardi, K. M. Morales, and K. A. Dooley, "Visible and infrared imaging spectroscopy of paintings and improved reflectography," *Heritage Science*, vol. 4, pp. 1–10, 2016.

[191] K. A. Dooley, S. Lomax, J. G. Zeibel, C. Miliani, P. Ricciardi, A. Hoenigswald, M. Loew, and J. K. Delaney, "Mapping of egg yolk and animal skin glue paint binders in Early Renaissance paintings using near infrared reflectance imaging spectroscopy," *Analyst*, vol. 138, no. 17, pp. 4838–4848, 2013.

[192] A. Deneckere, M. De Reu, M. Martens, K. De Coene, B. Vekemans, L. Vincze, P. De Maeyer, P. Vandenabeele, and L. Moens, "The use of a multi-method approach to identify the pigments in the 12th century manuscript Liber Floridus," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 80, no. 1, pp. 125–132, 2011.

[193] M. Aceto, A. Agostino, G. Fenoglio, M. Gulmini, V. Bianco, and E. Pellizzi, "Non invasive analysis of miniature paintings: proposal for an analytical protocol," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 91, pp. 352–359, 2012.

[194] S. Mosca, T. Frizzi, M. Pontone, R. Alberti, L. Bombelli, V. Capogrosso, A. Nevin, G. Valentini, and D. Comelli, "Identification of pigments in different layers of illuminated manuscripts by X-ray fluorescence mapping and Raman spectroscopy," *Microchemical Journal*, vol. 124, pp. 775–784, 2016.

[195] S. R. Amato, A. Burnstock, M. Cross, K. Janssens, F. Rosi, L. Cartechini, R. Fontana, A. D. Fovo, M. Paolantoni, C. Grazia, *et al.*, "Interpreting technical evidence from spectral imaging of paintings by édouard Manet in the Courtauld Gallery," *X-Ray Spectrometry*, vol. 48, no. 4, pp. 282–292, 2019.

[196] A. Dal Fovo, A. Mazzinghi, S. Omarini, E. Pampaloni, C. Ruberto, J. Striova, and R. Fontana, "Non-invasive mapping methods for pigments analysis of Roman mural paintings," *Journal of Cultural Heritage*, vol. 43, pp. 311–318, 2020.

[197] N. S. Daly, M. Sullivan, L. Lee, J. K. Delaney, and K. Trentelman, "Odilon Redon's noir drawings: characterization of materials and methods using noninvasive imaging and spectroscopies," *Heritage Science*, vol. 7, pp. 1–13, 2019.

[198] M. Thoury, J. K. Delaney, E. R. d. l. Rie, M. Palmer, K. Morales, and J. Krueger, "Near-infrared luminescence of cadmium pigments: in situ identification and mapping in paintings," *Applied Spectroscopy*, vol. 65, no. 8, pp. 939–951, 2011.

[199] J. K. Delaney, D. M. Conover, K. A. Dooley, L. Glinsman, K. Janssens, and M. Loew, "Integrated x-ray fluorescence and diffuse visible-to-near-infrared reflectance scanner for standoff elemental and molecular spectroscopic imaging of paints and works on paper," *Heritage Science*, vol. 6, pp. 1–12, 2018.

[200] A. Dal Fovo, J. Striova, E. Pampaloni, A. Fedele, M. M. Morita, D. Amaya, F. Grazzi, M. Cimò, C. Cirrincione, and R. Fontana, "Rubens' painting as inspiration of a later tapestry: Non-invasive analyses provide insight into artworks' history," *Microchemical Journal*, vol. 153, p. 104472, 2020.

[201] C. Montagner, M. Bacci, S. Bracci, R. Freeman, and M. Picollo, "Library of UV–Vis–NIR reflectance spectra of modern organic dyes from historic pattern-card coloured papers," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 79, no. 5, pp. 1669–1680, 2011.

[202] T. Cavaleri, P. Buscaglia, S. Migliorini, M. Nervo, G. Piccablotto, A. Piccirillo, M. Pisani, D. Puglisi, D. Vaudan, and M. Zucco, "Pictorial materials database: 1200 combinations of pigments, dyes, binders and varnishes designed as a tool

for heritage science and conservation," *Applied Physics A*, vol. 123, pp. 1–15, 2017.

[203] CHSOS, "Multispectral Imaging – Pigments Checker Standard." `https://chsopensource.org/multispectral-imaging-pigments-checker-standard/`, 2024. Accessed: January 30, 2025.

[204] F. Lombardi and S. Marinai, "Deep learning for historical document analysis and recognition—a survey," *Journal of Imaging*, vol. 6, no. 10, p. 110, 2020.

[205] K. Nikolaidou, M. Seuret, H. Mokayed, and M. Liwicki, "A survey of historical document image datasets," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 25, no. 4, pp. 305–338, 2022.

[206] E. Granell, V. Romero, J. R. Prieto, J. Andrés, L. Quirós, J. A. Sánchez, and E. Vidal, "Processing a large collection of historical tabular images," *Pattern Recognition Letters*, vol. 170, pp. 9–16, 2023.

[207] I. Pratikakis, K. Zagoris, X. Karagiannis, L. Tsochatzidis, T. Mondal, and I. Marthot-Santaniello, "Competition on document image binarization (DIBCO 2019)," in *2019 15th International Conference on Document Analysis and Recognition*, pp. 1547–1556, IEEE, 2019.

[208] I. Pratikakis, K. Zagoris, G. Barlas, and B. Gatos, "ICFHR2018 competition on handwritten document image binarization contest (H-DIBCO 2018)," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 1–5, IEEE, 2018.

[209] N. Serrano, F. Castro, and A. Juan, "The RODRIGO database.," in *LREC*, pp. 19–21, 2010.

[210] A. Alaei, P. Nagabhushan, and U. Pal, "A new dataset of Persian handwritten documents and its segmentation," in *2011 7th Iranian conference on machine vision and image processing*, pp. 1–5, IEEE, 2011.

[211] A. Abdelhaleem, A. Droby, A. Asi, M. Kassis, R. Al Asam, and J. El-sanaa, "Wahd: a database for writer identification of arabic historical documents," in *2017 1st International workshop on arabic script analysis and recognition (ASAR)*, pp. 64–68, IEEE, 2017.

[212] A. Shahkolaei, A. Beghdadi, S. Al-Máadeed, and M. Cheriet, "MHDID: a multi-distortion historical document image database," in *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)*, pp. 156–160, IEEE, 2018.

[213] M. Lettner and R. Sablatnig, "Higher order mrf for foreground-background separation in multi-spectral images of historical manuscripts," in *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*, pp. 317–324, 2010.

[214] Y. E. Salehani, E. Arabnejad, A. Rahiche, A. Bakhta, and M. Cheriet, "MSdB-NMF: MultiSpectral document image binarization framework via non-negative matrix factorization approach," *IEEE Transactions on Image Processing*, vol. 29, pp. 9099–9112, 2020.

[215] R. Hedjam and M. Cheriet, "Novel data representation for text extraction from multispectral historical document images," in *2011 International Conference on Document Analysis and Recognition*, pp. 172–176, IEEE, 2011.

[216] R. Hedjam, H. Z. Nafchi, R. F. Moghaddam, M. Kalacska, and M. Cheriet, "Icdar 2015 contest on multispectral text extraction (ms-tex 2015)," in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1181–1185, IEEE, 2015.

[217] N. Mitianoudis and N. Papamarkos, "Multi-spectral document image binarization using image fusion and background subtraction techniques," in *2014 IEEE International Conference on Image Processing (ICIP)*, pp. 5172–5176, IEEE, 2014.

[218] M. Diem, F. Hollaus, and R. Sablatnig, "Msio: Multispectral document image binarization," in *2016 12th IAPR Workshop on Document Analysis Systems (DAS)*, pp. 84–89, IEEE, 2016.

[219] R. F. Moghaddam and M. Cheriet, "A multi-scale framework for adaptive binarization of degraded document images," *Pattern Recognition*, vol. 43, no. 6, pp. 2186–2198, 2010.

[220] F. Hollaus, M. Diem, and R. Sablatnig, "Binarization of multispectral document images," in *Computer Analysis of Images and Patterns: 16th International Conference, CAIP 2015, Valletta, Malta, September 2-4, 2015, Proceedings, Part II 16*, pp. 109–120, Springer, 2015.

[221] Z. Khan, F. Shafait, and A. Mian, "Automatic ink mismatch detection for forensic document analysis," *Pattern Recognition*, vol. 48, no. 11, pp. 3615–3626, 2015.

[222] A. U. Islam, M. J. Khan, M. Asad, H. A. Khan, and K. Khurshid, "iVision HHID: handwritten hyperspectral images dataset for benchmarking hyperspectral imaging-based document forensic analysis," *Data in Brief*, vol. 41, p. 107964, 2022.

[223] A. Morales, M. A. Ferrer, M. Diaz-Cabrera, C. Carmona, and G. L. Thomas, "The use of hyperspectral analysis for ink identification in handwritten documents," in *2014 International Carnahan Conference on Security Technology (ICCST)*, pp. 1–5, IEEE, 2014.

[224] A. B. López-Baldomero, M. Martínez-Domingo, E. M. Valero, R. Fernández-Gualda, A. López-Montes, R. Blanc-García, and T. Espejo, "Selection of optimal spectral metrics for classification of inks in historical documents using hyperspectral imaging data," in *Optics for Arts, Architecture, and Archaeology (O3A) IX*, vol. 12620, pp. 99–111, SPIE, 2023.

[225] "Color Imaging Lab - HYPERDOC project." `https://colorimaginglab.ugr.es/pages/hyperdoc/project`, 2025. Accessed: January 30, 2025.

[226] A. B. López-Baldomero, J. L. Nieves, F. Moronta-Montero, M. A. Martínez-Domingo, R. Fernández-Gualda, J. Hernández-Andrés, A. S. Reichert, A. López-Montes, T. Espejo, J. Romero, and E. M. Valero, "Google Drive - HYPERDOC database.." `https://drive.google.com/drive/folders/1ZmhX3NZhqKWVTvmGmM_XHY18iDoI2NGP?usp=sharing`, 2025.

[227] D. M. Goltz, "A review of instrumental approaches for studying historical inks," *Analytical letters*, vol. 45, no. 4, pp. 314–329, 2012.

[228] G. M. C. Zamorano, "The presence of iron in inks used in valencian manuscripts from the 13th to 17th century," *Microchemical Journal*, vol. 143, pp. 484–492, 2018.

[229] A. B. López-Baldomero, M. Buzzelli, F. Moronta-Montero, M. Á. Martínez-Domingo, and E. M. Valero, "Ink classification in historical documents using hyperspectral imaging and machine learning methods," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 335, p. 125916, 2025.

[230] A. B. López-Baldomero, E. Valero, A. Reichert, F. Moronta-Montero, M. Martínez-Domingo, and A. López-Montes, "Hyperspectral database of synthetic historical inks," in *Archiving Conference*, vol. 21, pp. 11–16, Society for Imaging Science and Technology, 2024.

[231] F. Moronta-Montero, R. Gualda, A. B. López-Baldomero, M. Buzzelli, M. Martínez-Domingo, and E. Valero, "Evaluation of binarization methods for hyperspectral samples of 16th and 17th century family trees," *Archiving Conference*, vol. 21, pp. 94–100, 04 2024.

[232] M. Buzzelli, F. Moronta-Montero, R. Fernández-Gualda, A. B. López-Baldomero, J. L. Nieves, and E. M. Valero, "Handwritten ink segmentation algorithms for hyperspectral images of historical documents." Under review, 2024.

[233] A. B. López-Baldomero, E. M. Valero, M. A. Martínez-Domingo, A. S. Reichert, and A. López-Montes, "Spectral unmixing to identify historical inks." Colour and Visual Computing Symposium (CVCS), Gjøvik, 2024.

[234] A. S. Reichert, A. B. López-Baldomero, F. J. Collado-Montero, and A. López-Montes, "Study of traditional writing inks: creation of standard samples and colorimetric analysis," in *TechnoHeritage2024. Book of Abstracts*, p. 118, Universidade de Vigo, 2024.

[235] E. M. Valero, A. B. López-Baldomero, A. S. Reichert, and A. López-Montes, "Hyperspectral insights into iron gall ink aging," in *TechnoHeritage2024. Book of Abstracts*, p. 182, Universidade de Vigo, 2024.

[236] R. J. Díaz, R. Cordoba, H. Grigoryan, M. Vieira, M. Melo, P. Nabais, V. Otero, N. Teixeira, S. Fani, and H. Al-Abbady, "The making of black inks in an arabic treatise by al-Qalalūsī dated from the 13th c.: reproduction and characterisation of iron-gall ink recipes," *Heritage Science*, vol. 11, pp. 1–14, 2023.

[237] G. M. Contreras Zamorano, *La tinta de escritura en los manuscritos de archivo valencianos, 1250-1600. Análisis, identificación de componentes y valoración de su estado de conservación*. PhD thesis, Universitat de València, 2015.

[238] D. Conn and D. Walus, "Paper objects and books," *The Preservation Management Handbook: A 21st-Century Guide for Libraries, Archives, and Museums*, pp. 147–181, 2014.

[239] T. Espejo, I. Lazarova, and M. Cano, "La colección de manuscritos árabes del Archivo Histórico Provincial de Granada. primeros apuntes sobre su caracterización," in *VIII Congreso Nacional de Historia del Papel en España: Actas*, pp. 33–44, 2008.

[240] T. E. Arias, I. L. Stoytcheva, D. C. García, A. D. Benito, and A. J. de Haro, "Caracterización material y proceso de conservación de la colección de documentos árabes manuscritos del Archivo Histórico Provincial de Granada," *Al-Qanṭara*, vol. 32, no. 2, pp. 519–532, 2011.

[241] M. L. de Guevara *et al.*, "Pleitos de Hidalguía. Extracto de sus expedientes que se conservan en el Archivo de la Real Chancillería de Granada correspondiente a la primera parte del reinado de Felipe II (1556-1570): en cuatro volúmenes," *Hidalgos: la revista de la Real Asociación de Hidalgos de España*, vol. 1, no. 565, pp. 94–95, 2021.

[242] A. Duran, A. López-Montes, J. Castaing, and T. Espejo, "Analysis of a royal 15th century illuminated parchment using a portable XRF–XRD system and micro-invasive techniques," *Journal of archaeological science*, vol. 45, pp. 52–58, 2014.

[243] "Color Imaging Lab - Hyperspectral capture of illuminated manuscript @ Royal Chancery Archive of Granada, Spain." `https://www.youtube.com/watch?v=lIwNyjBeKmQ`, 2025. Accessed: January 30, 2025.

[244] A. B. López-Baldomero, F. Moronta-Montero, M. A. Martínez-Domingo, and E. M. Valero, "HYPERDOC_Database_code." `https://github.com/anabelenlb/HYPERDOC_Database_code`, 2025. Accessed: January 30, 2025.

[245] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9*, pp. 404–417, Springer, 2006.

[246] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[247] T.-C. Lee, R. L. Kashyap, and C.-N. Chu, "Building skeleton models via 3-D medial surface axis thinning algorithms," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 6, pp. 462–478, 1994.

[248] K. Ntirogiannis, B. Gatos, and I. Pratikakis, "An objective evaluation methodology for document image binarization techniques," in *2008 The Eighth IAPR International Workshop on Document Analysis Systems*, pp. 217–224, IEEE, 2008.

[249] N. Otsu *et al.*, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[250] W. Niblack, *An introduction to digital image processing*. Strandberg Publishing Company, 1985.

[251] C. Wolf, J.-M. Jolion, and F. Chassaing, "Text localization, enhancement and binarization in multimedia documents," in *2002 International Conference on Pattern Recognition*, vol. 2, pp. 1037–1040, IEEE, 2002.

[252] D. Bradley and G. Roth, "Adaptive thresholding using the integral image," *J. Graphics Tools*, vol. 12, pp. 13–21, 01 2007.

[253] N. R. Howe, "Document binarization with automatic parameter tuning," *International journal on document analysis and recognition (ijdar)*, vol. 16, pp. 247–258, 2013.

[254] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern recognition*, vol. 33, no. 2, pp. 225–236, 2000.

[255] F. Grillini, L. Aksas, P.-J. Lapray, A. Foulonneau, J.-B. Thomas, S. George, and L. Bigué, "Relationship between reflectance and degree of polarization in the

VNIR-SWIR: A case study on art paintings with polarimetric reflectance imaging spectroscopy," *Plos one*, vol. 19, no. 5, p. e0303018, 2024.

[256] CHSOS, "Reflectance Spectra Database for the new Pigments Checker "Modern & Contemporary Art"." `https://chsopensource.org/pigments-checker/`, 2024. Accessed: January 30, 2025.

[257] H. Deborah, "Hyperspectral pigments." `https://hyppigments.streamlit.app/`, 2022. Accessed: January 30, 2025.

[258] Z. Khan, F. Shafait, and A. Mian, "Hyperspectral imaging for ink mismatch detection," in *2013 12th international conference on document analysis and recognition*, pp. 877–881, IEEE, 2013.

[259] C. S. Silva, M. F. Pimentel, R. S. Honorato, C. Pasquini, J. M. Prats-Montalbán, and A. Ferrer, "Near infrared hyperspectral imaging for forensic analysis of document forgery," *Analyst*, vol. 139, no. 20, pp. 5176–5184, 2014.

[260] M. Agarla, S. Bianco, L. Celona, R. Schettini, and M. Tchobanou, "An analysis of spectral similarity measures," in *Color and Imaging Conference*, vol. 29, pp. 300–305, Society for Imaging Science and Technology, 2021.

[261] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote sensing of environment*, vol. 37, no. 1, pp. 35–46, 1991.

[262] C.-I. Chang, "Spectral information divergence for hyperspectral image analysis," in *IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No. 99CH36293)*, vol. 1, pp. 509–511, IEEE, 1999.

[263] Y. Du, C.-I. Chang, H. Ren, C.-C. Chang, J. O. Jensen, and F. M. D'Amico, "New hyperspectral discrimination measure for spectral characterization," *Optical engineering*, vol. 43, no. 8, pp. 1777–1786, 2004.

[264] R. R. Nidamanuri and B. Zbell, "Normalized spectral similarity score (NS^3) as an efficient spectral library searching method for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 1, pp. 226–240, 2010.

[265] M. Dabboor, S. Howell, M. Shokr, and J. Yackel, "The Jeffries–Matusita distance for the case of complex wishart distribution as a separability criterion for fully polarimetric SAR data," *International journal of remote sensing*, vol. 35, no. 19, pp. 6859–6873, 2014.

[266] L. Tan *et al.*, "A study on the application of SAM classification algorithm in seal of calligraphy and painting based on hyperspectral technology," in *2016*

4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA), pp. 415–418, IEEE, 2016.

[267] L. Krauz, P. Páta, and J. Kaiser, "Assessing the spectral characteristics of dye-and pigment-based inkjet prints by vnir hyperspectral imaging," *Sensors*, vol. 22, no. 2, p. 603, 2022.

[268] L. Bruzzone, F. Roli, and S. B. Serpico, "An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 6, pp. 1318–1321, 1995.

[269] S. Padma and S. Sanjeevi, "Jeffries Matusita-Spectral Angle Mapper (JM-SAM) spectral matching for species level mapping at Bhitarkanika, Muthupet and Pichavaram mangroves," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, pp. 1403–1411, 2014.

[270] F. Van der Meer, "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery," *International journal of applied earth observation and geoinformation*, vol. 8, no. 1, pp. 3–17, 2006.

[271] C.-I. Chang, "An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis," *IEEE Transactions on information theory*, vol. 46, no. 5, pp. 1927–1932, 2000.

[272] B. Halder and A. C. Mondal, "Modelling of ink-colour degradation on old printed documents," *International Journal of Computational Vision and Robotics*, vol. 7, no. 4, pp. 398–410, 2017.

[273] R. T. Matenda, D. Rip, J. A. F. Pierna, V. Baeten, and P. J. Williams, "Differentiation of listeria monocytogenes serotypes using near infrared hyperspectral imaging," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 320, p. 124579, 2024.

[274] Y. Shao, S. Ji, Y. Shi, G. Xuan, H. Jia, X. Guan, and L. Chen, "Growth period determination and color coordinates visual analysis of tomato using hyperspectral imaging technology," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, p. 124538, 2024.

[275] J. Yang, L. Sun, W. Xing, G. Feng, H. Bai, and J. Wang, "Hyperspectral prediction of sugarbeet seed germination based on gauss kernel svm," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 253, p. 119585, 2021.

[276] E. Catelli, L. L. Randeberg, B. K. Alsberg, K. F. Gebremariam, and S. Bracci, "An explorative chemometric approach applied to hyperspectral images for the study of illuminated manuscripts," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 177, pp. 69–78, 2017.

[277] G. Bonifazi, G. Capobianco, P. Cucuzza, and S. Serranti, "Hyperspectral imaging coupled with data fusion for plastic packaging waste recycling," in *SPIE Future Sensing Technologies 2023*, vol. 12327, pp. 104–116, SPIE, 2023.

[278] A. Smolinska, J. Engel, E. Szymanska, L. Buydens, and L. Blanchet, "General framing of low-, mid-, and high-level data fusion with examples in the life sciences," in *Data Handling in Science and Technology*, vol. 31, pp. 51–79, Elsevier, 2019.

[279] J. Chen, C. Fu, and T. Pan, "Modeling method and miniaturized wavelength strategy for near-infrared spectroscopic discriminant analysis of soy sauce brand identification," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 277, p. 121291, 2022.

[280] R. Ullah, I. Rehan, and S. Khan, "Utilizing machine learning algorithms for precise discrimination of glycosuria in fluorescence spectroscopic data," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, p. 124582, 2024.

[281] P. Xanthopoulos, P. M. Pardalos, T. B. Trafalis, P. Xanthopoulos, P. M. Pardalos, and T. B. Trafalis, "Linear discriminant analysis," *Robust data mining*, pp. 27–33, 2013.

[282] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 801–818, 2018.

[283] M. Iman, H. R. Arabnia, and K. Rasheed, "A review of deep transfer learning and recent advancements," *Technologies*, vol. 11, no. 2, p. 40, 2023.

[284] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pp. 740–755, Springer, 2014.

[285] The MathWorks, Inc., "fitcsvm." `https://es.mathworks.com/help/stats/fitcsvm.html`, 2024. [Online; accessed 30-Jul-2024].

[286] V. Corregidor, R. Viegas, L. M. Ferreira, and L. C. Alves, "Study of iron gall inks, ingredients and paper composition using non-destructive techniques," *Heritage*, vol. 2, no. 4, pp. 2691–2703, 2019.

[287] M. Faries, "Analytical capabilities of infrared reflectography: an art historian's perspective," *Scientific examination of art: Modern techniques in conservation and analysis*, pp. 87–104, 2005.

[288] V. Chelladurai and D. Jayas, "Near-infrared imaging and spectroscopy," in *Imaging with Electromagnetic Spectrum: Applications in Food and Agriculture*, pp. 87–127, Springer, 2014.

[289] D. Mazzini, M. Buzzelli, D. P. Pauy, and R. Schettini, "A CNN architecture for efficient semantic segmentation of street scenes," in *2018 IEEE 8th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*, pp. 1–6, IEEE, 2018.

[290] B. Havlınová, D. Babiaková, V. Brezová, M. Ďurovič, M. Novotná, and F. Belányi, "The stability of offset inks on paper upon ageing," *Dyes and pigments*, vol. 54, no. 2, pp. 173–188, 2002.

[291] P. Ricciardi, J. K. Delaney, M. Facini, J. G. Zeibel, M. Picollo, S. Lomax, and M. Loew, "Near infrared reflectance imaging spectroscopy to map paint binders in situ on illuminated manuscripts," *Angewandte Chemie International Edition*, vol. 51, no. 23, pp. 5607–5610, 2012.

[292] F. Grillini, J. B. Thomas, and S. George, "VisNIR pigment mapping and rerendering of an experimental painting," *Journal of the International Colour Association*, 2021.

[293] R. Blanc, E. Manzano, A. López-Montes, N. Domínguez-Gasca, and J. L. Vílchez, "Non-invasive study of the pigments of a painting on copper with the inscription "Boceto di Pablo Veronese" on the back," *Heritage*, vol. 6, no. 6, pp. 4787–4801, 2023.

[294] L. F. López, I. C. Blanco, M. F. S. Martín, M. L. V. de Ágredos Pasqual, L. Carlyle, and J. Wadum, *Paintings on Copper and other metal plates: Production, degradation and conservation issues*. Colección UPV Scientia, 2017.

[295] I. Horovitz, "The materials and techniques of european paintings on copper supports," *Copper as canvas: two centuries of masterpiece paintings on copper*, pp. 1575–1775, 1999.

[296] F. Pacheco, *Arte de la pintura: su antigüedad y grandezas*, vol. 1. Galiano, 1866.

[297] D. Vega, I. Pombo Cardoso, and L. Carlyle, "Pintura sobre cobre: investigación sobre materiales y técnicas de aplicación de la capa de preparación a través de los tratados tradicionales y estudio analítico de dos obras atribuidas a las escuelas portuguesa y flamenca.," *Conservar Património*, no. 27, 2018.

[298] J. Romero, A. Garcıa-Beltrán, and J. Hernández-Andrés, "Linear bases for representation of natural and artificial illuminants," *JOSA A*, vol. 14, no. 5, pp. 1007–1014, 1997.

[299] F. A. Potra and S. J. Wright, "Interior-point methods," *Journal of computational and applied mathematics*, vol. 124, no. 1-2, pp. 281–302, 2000.

[300] S. A. Burns, "Subtractive color mixture computation," *arXiv preprint arXiv:1710.06364*, 2017.

[301] F. H. Imai, M. R. Rosen, and R. S. Berns, "Comparative study of metrics for spectral match quality," in *Conference on colour in graphics, imaging, and vision*, vol. 1, pp. 492–496, Society of Imaging Science and Technology, 2002.

[302] J. Cai, H. Chatoux, C. Boust, and A. Mansouri, "Extending the unmixing methods to multispectral images," in *Proceedings of the Color and Imaging Conference*, pp. 311–316, Society for Imaging Science and Technology, 2021.

[303] C.-I. Chang and A. Plaza, "A fast iterative algorithm for implementation of pixel purity index," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 1, pp. 63–67, 2006.

[304] L. Simonot and M. Hébert, "Between additive and subtractive color mixings: intermediate mixing models," *Journal of the Optical Society of America A*, vol. 31, no. 1, pp. 58–66, 2013.

[305] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Summaries of the fifth annual JPL airborne earth science workshop. Volume 1: AVIRIS workshop*, 1995.

[306] J. W. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," in *JPL, Summaries of the 4th Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*, 1993.

[307] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tourneret, A. Zare, and C. Jutten, "Spectral variability in hyperspectral data unmixing: A comprehensive review," *IEEE geoscience and remote sensing magazine*, vol. 9, no. 4, pp. 223–270, 2021.

[308] L. Ortiz-Herrero, I. Cardaba, L. Bartolomé, M. Alonso, and M. Maguregui, "Extension study of a statistical age prediction model for acrylic paints," *Polymer Degradation and Stability*, vol. 179, p. 109263, 2020.

[309] R. A. Waltz, J. L. Morales, J. Nocedal, and D. Orban, "An interior algorithm for nonlinear optimization that combines line search and trust region steps," *Mathematical programming*, vol. 107, no. 3, pp. 391–408, 2006.

[310] F. Grillini, J.-B. Thomas, and S. George, "Logistic splicing correction for VNIR–SWIR reflectance imaging spectroscopy," *Optics Letters*, vol. 48, no. 2, pp. 403–406, 2023.

[311] Ideal, newspaper, "La UGR crea la mayor colección de datos de imagen espectral de documentos históricos del mundo." `https://www.ideal.es/miugr/`

ugr-crea-mayor-coleccion-datos-imagen-espectral-20250217112448-nt.html, 2025. [Online; accessed 11-Mar-2025].

[312] Granada Hoy, newspaper, "La Universidad de Granada crea la mayor colección de datos de imagen espectral de documentos históricos del mundo." https://www.granadahoy.com/granada/universidad-granada-mayor-coleccion-mundo-documentos-historicos_0_2003352354.html, 2025. [Online; accessed 11-Mar-2025].

[313] Canal UGR, newspaper, "La UGR crea Hyperdoc, la mayor colección de datos de imagen espectral de documentos históricos del mundo." https://canal.ugr.es/noticia/la-ugr-crea-hyperdoc/, 2025. [Online; accessed 11-Mar-2025].

[314] Color Imaging Lab UGR, YouTube, "Hyperdoc database at regional public TV CanalSur." https://www.youtube.com/watch?v=Uv13fnfXwRk, 2025. [Online; accessed 11-Mar-2025].

List of Figures

List of Tables

Scientific production

JOURNAL ARTICLES INCLUDED IN THIS PHD THESIS

A. B. López-Baldomero, M. Buzzelli, F. Moronta-Montero, M. Á. Martínez-Domingo, and E. M. Valero, "Ink classification in historical documents using hyperspectral imaging and machine learning methods," *Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy*, vol. 335, pp. 125916, 2025.

E. M. Valero, M. A. Martínez-Domingo, **A. B. López-Baldomero**, A. López-Montes, D. Abad-Muñoz, and J. L. Vílchez-Quero, "Unmixing and pigment identification using visible and short-wavelength infrared: Reflectance vs logarithm reflectance hyperspaces," *Journal of Cultural Heritage*, vol. 64, pp. 290–300, 2023.

A. B. López-Baldomero, J. L. Nieves, F. Moronta-Montero, M. A. Martínez-Domingo, R. Fernández-Gualda, J. Hernández-Andrés, A. S. Reichert, A. López-

Montes, T. Espejo, J. Romero, and E. M. Valero, "HYPERDOC - HYPERspectral database of historical DOCuments and mock-ups from 400 to 1700 nm" (Submitted to *Scientific Data*, 2025).

OTHER JOURNAL ARTICLES

M. Á. Martínez-Domingo, **A. B. López-Baldomero**, M. Tejada-Casado, M. Melgosa, and F. J. Collado-Montero, "Colorimetric Evaluation of a Reintegration via Spectral Imaging—Case Study: Nasrid Tiling Panel from the Alhambra of Granada (Spain)", *Sensors*, 24(12), 3872, 2024.

S. Yang, H. Qin, Y. Dai, X. Yan, and **A. B. López-Baldomero**, "Temperature distribution inversion in infrared multispectral imaging based on ensemble network", *Optics Letters*, 49(18), 5163-5166, 2024.

A. B. López-Baldomero, M. Rubino, C. Ortiz, and C. Salas, "New software for comparing the color gamuts generated by printing technologies", *Optik*, 261, 169194, 2022.

CONFERENCES AND SCIENTIFIC MEETINGS

A. B. López-Baldomero, M. Martínez-Domingo, E. M. Valero, R. Fernández-Gualda, A. López-Montes, R. Blanc-García, and T. Espejo, "Selection of optimal spectral metrics for classification of inks in historical documents using hyperspectral imaging data," in *Optics for Arts, Architecture, and Archaeology (O3A) IX*, vol. 12620, pp. 99–111, SPIE, 2023.

A. B. López-Baldomero, M. A. Martínez-Domingo, J. Hernández-Andrés, R. Blanc, J. Vilchez-Quero, A. López-Montes, and E. M. Valero, "Endmember extraction for pigment identification pre-and post-intervention: A case study from a XVIth

century copper plate painting," in *Archiving Conference*, vol. 20, pp. 198–203, 2023.

A. B. López-Baldomero, E. Valero, A. Reichert, F. Moronta-Montero, M. Martínez-Domingo, and A. López-Montes, "Hyperspectral database of synthetic historical inks," in *Archiving Conference*, vol. 21, pp. 11–16, 2024.

A. B. López-Baldomero, E. M. Valero, M. A. Martínez-Domingo, A. S. Reichert, and A. López-Montes, "Spectral unmixing to identify historical inks," in *Colour and Visual Computing Symposium (CVCS)*, Gjøvik (Norway), 2024.

F. Moronta-Montero, R. Fernández-Gualda, **A. B. López-Baldomero**, M. Buzzelli, M. A. Martínez-Domingo, and E. M. Valero, "Evaluation of Binarization Methods for Hyperspectral Samples of 16th and 17th Century Family Trees", in *Archiving Conference*, vol. 21, pp. 94-100, 2024.

A. B. López-Baldomero, E. M. Valero, M. A. Martínez-Domingo, A. D. Abad-Muñoz, and F. Moronta-Montero, "Identificando la paleta del artista a partir de la imagen espectral," in *V Congreso/VII Jornadas de Investigadores/as en Formación*, Granada (Spain), 2024.

F. Moronta-Montero, **A. B. López-Baldomero**, R. Fernández Gualda, M. A. Martínez-Domingo, and E. M. Valero, "Clasificación de tintas en documentos históricos mediante imagen hiperspectral," in *V Congreso/VII Jornadas de Investigadores/as en Formación*, Granada (Spain), 2024.

E. M. Valero, **A. B. López-Baldomero**, A. S. Reichert, and A. López-Montes, "Hyperspectral insights into iron gall ink aging," in *TechnoHeritage 2024*, Santiago de Compostela (Spain), 2024.

A. S. Reichert, **A. B. López-Baldomero**, F. J. Collado-Montero, and A. López-

Montes, "Study of traditional writing inks: creation of standard samples and colorimetric analysis," in *TechnoHeritage 2024*, Santiago de Compostela (Spain), 2024.

A. B. López-Baldomero, E. M. Valero, M. A. Martínez-Domingo, A. A. S. Reichert, and A. López-Montes, "Identificación de mezclas de tintas históricas mediante técnicas de unmixing espectral," in *XIV Reunión Nacional de Óptica (RNO)*, Murcia (Spain), 2024.

RESEARCH STAY

Colourlab, Department of Computer Science, Norwegian University of Science and Technology (NTNU), Gjøvik. **3 months.** Supervisor: Prof. Sony George.

PHD SCHOLARSHIPS

Predoctoral Fellowship (FPU) funded by the Ministry of Universities, Government of Spain (FPU2020-05532). Duration: 01/10/2022 – 18/05/2026.

Mobility Grant for Short Research Stays under the FPU subprogram, funded by the Ministry of Science, Innovation, and Universities. Duration: 15/08/2024 – 13/11/2024.

PARTICIPATION IN RESEARCH PROJECTS

"Identificación automática de materiales en documentos gráficos de interés histórico-artístico mediante imagen hiperespectral (HYPERDOC)" (PID2021-124446NB-I00). Funded by MICIU/AEI /10.13039/501100011033 and by ERDF, EU. University of Granada. P.I. Eva M. Valero Benito. 30/08/2022 - 30/08/2025

"El museo de lo invisible. Técnicas de imagen espectral para la recuperación digital de patrimonio artístico, arqueológico y documental deteriorado" (N607D

2023/13). Funded by Agencia Gallega de Innovación (GAIN), Xunta de Galicia, Proyectos de Excelencia, Modalidad C: Consolidación y estructuración de unidades de investigación competitivas. Instituto de Ciencias del Patrimonio (INCIPIT). P.I. Lucía Pereira Pardo. 15/11/2023 - 31/12/2027

"Autentificación de mieles amparadas bajo el sello de calidad Denominación de Origen Protegida (DOP) mediante análisis de imágenes hiperespectrales" (PPJIB-2024-25). Funded by University of Granada, Plan Propio de Investigación y Transferencia, P20 Proyectos de Investigación Precompetitivos para Jóvenes Investigadores. University of Granada. P.I. Miriam Medina García and **Ana Belén López Baldomero**. 01/01/2025 - 31/12/2025

TEACHING ACTIVITIES

Lectures in the Department of Optics at the University of Granada, teaching Physiological Optics I, Physiological Optics II, and Optometric Instrumentation in the Bachelor's Degree in Optics and Optometry, and Optics II in the Bachelor's Degree in Physics (180 hours).

Practical seminars in the Erasmus+ Computational Colour and Spectral Imaging (COSI) Master's Program and the Master's in Science and Technology for Architectural Heritage at the University of Granada.