Chemometrics and Intelligent Laboratory Systems 248 (2024) 105116

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics

Machine learning regression algorithms for generating chemical element maps from X-ray fluorescence data of paintings

Juan Ruiz de Miras [a,*], María José Gacto [a], María Rosario Blanc [b], Germán Arroyo [a], Luis López [a], Juan Carlos Torres [a], Domingo Martín [a]

[a] Software Engineering Department, University of Granada, Granada, Spain
[b] Department of Analytical Chemistry, University of Granada, Spain

ARTICLE INFO

Keywords:
Machine learning
Random forest
Regression problem
X-ray fluorescence
Chemical element maps

ABSTRACT

Generating chemical element maps of paintings from X-ray fluorescence (XRF) data is a very valuable tool for the scientific community of conservators and art historians. Hand-held XRF scanners are cheap and easily portable but their use provides scans with a few data, so additional analytical tools are needed to obtain reliable chemical element maps from them. Recently, the software tool *SmART_Scan* was released, which uses an algorithm based on the minimum hypercube distance (MHD) to compute this kind of maps. In this paper, we propose a new methodology to address this problem by using machine learning algorithms for regression as alternative and more accurate techniques than MHD. We tested MHD versus eight machine learning regression algorithms on two paintings with different features. Our results showed that machine learning algorithms Random Forest and kNN significantly outperformed MHD in Mean Squared Error (MSE) and coefficient of determination (R^2) for all the experiments. When using experts' data and a hold-out validation, kNN was the best-ranked algorithm. Random Forest was the best-ranked algorithm when cross-validation was used. We did not find significant differences in average MSE nor in R^2 between kNN and Random Forest, so we can conclude that Random Forest is the best-suited algorithm for computing chemical element maps of paintings from XRF data.

1. Introduction

X-ray fluorescence (XRF) scanners are very valuable tools for analyzing the distribution of pigments on paintings, among many other applications [1]. These scanners are very useful for the scientific community of conservators and art historians because analyzing the pigments used in a painting allow them to date the work, identify previous restorations and perform conservation techniques [2]. An XRF scan on a certain point of a painting returns an energy spectrum whose peaks correspond to the XRF emissions of chemical elements. As the characteristic energy of each chemical element is known, it is possible to identify the elements present in that point from the XRF spectrum [2].

XRF scanners can be grouped into two main families: MacroXRF and hand-held scanners. MacroXRF scanners produce very precise results allowing the users to scan thousands of samples on the painting, but they are much more expensive and difficult to transport than hand-held portable scanners. These are the reasons why the development of efficient and accurate data analysis tools for those less expensive and less complex hand-held scanners is an important focus of study nowadays [1].

In this line of work, Martin-Ramos et al. presented *SmART_Scan* [3], a computer program for generating chemical element maps showing the distribution of chemical elements in the painting. These maps are generated using as inputs an RGB image of the painting and a limited number of XRF scanned points selected by an expert. SmART_Scan obtains the maps by combining the data through a method called Minimum Hypercube Distance (MHD). The short processing time and the quality of the maps generated position SmART_Scan as a valuable alternative approach to analyzing data obtained with MacroXRF scanners. To the best of our knowledge, MHD is the only published method for obtaining chemical element maps from scattered XRF data.

The MHD algorithm in SmART_Scan has been successfully used in several recent studies such as [4], where it was used to analyze two mural paintings located in the archaeological site of Pompeii (Italy); [5],

* Corresponding author. Software Engineering Department, Periodista Manuel Saucedo Aranda s/n, 18071, Granada, Spain.
E-mail addresses: demiras@ugr.es (J. Ruiz de Miras), mjgacto@ugr.es (M.J. Gacto), mrblanc@ugr.es (M.R. Blanc), arroyo@ugr.es (G. Arroyo), luislopez@ugr.es (L. López), jctorres@ugr.es (J.C. Torres), dmartin@ugr.es (D. Martín).

https://doi.org/10.1016/j.chemolab.2024.105116
Received 13 August 2023; Received in revised form 8 November 2023; Accepted 16 March 2024
Available online 25 March 2024

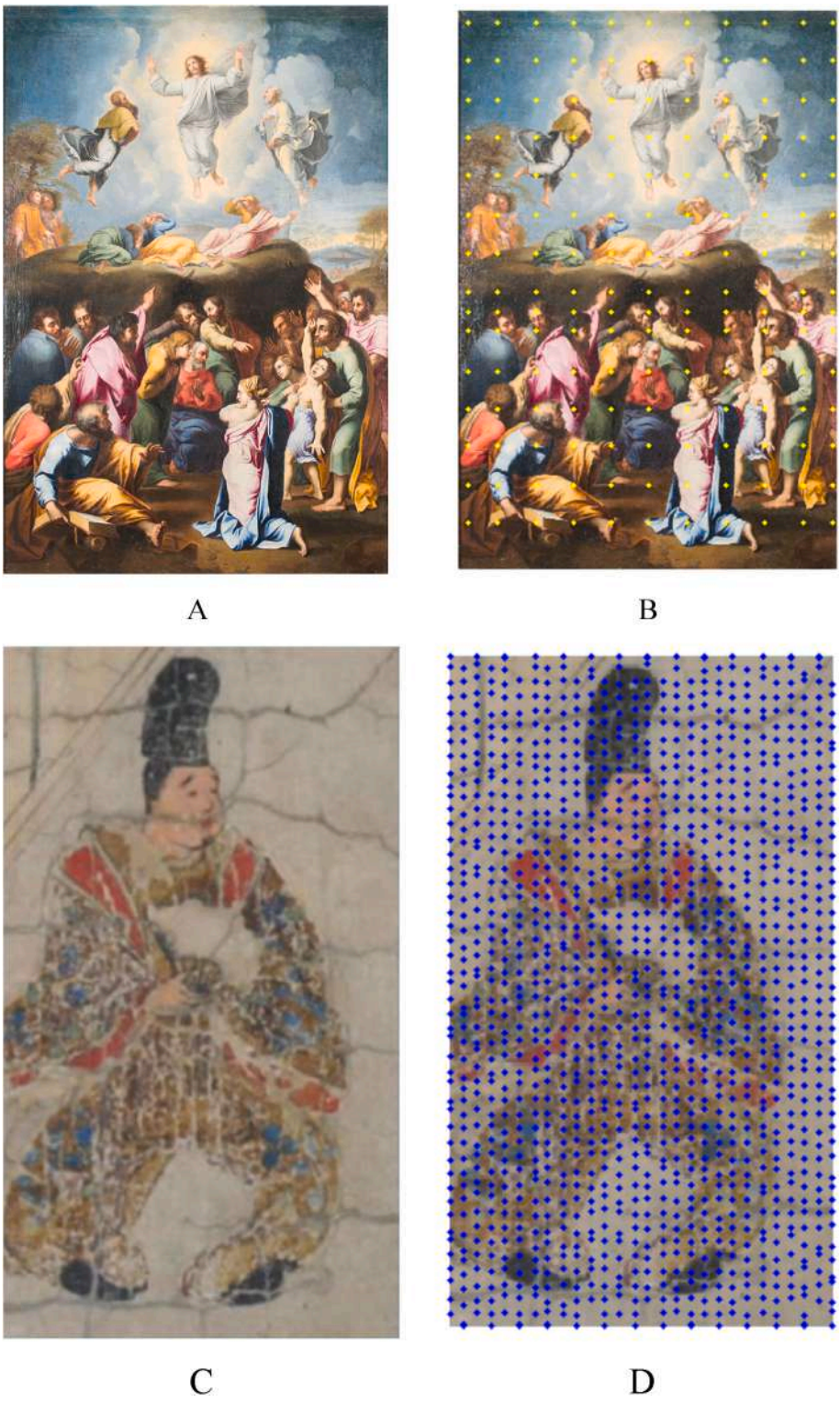

Fig. 1. Paintings used to test the accuracy of the algorithms: (A) The Transfiguration; (B) Locations of the 165 scanned points for The Transfiguration; (C) Man, a piece of the scroll The Miraculous Interventions of Jizō Bosatsu; and (D) Locations of the 1314 scanned points for Man.

where the pigments and materials in a copy of The Transfiguration painting by Raphael were characterized; and [6,7], two studies on the constituent materials and the execution techniques of scenes painted on antique vases by the Phiale Painter and from the Centuripe area, respectively.

Recently, several authors applied classification algorithms based on convolutional neural networks (CNN) for identifying pigments in paintings [8–11], which allowed them to predict the class of pigment. The algorithms described in these works generated a pigment map for the painting from the input data (XRF data or hyperspectral images) once the CNN has been trained with other paintings, preferably from the same epoch or author. However, these machine learning classification algorithms cannot be applied for extracting chemical element maps because they obtain a single label with the pigment in each point but not

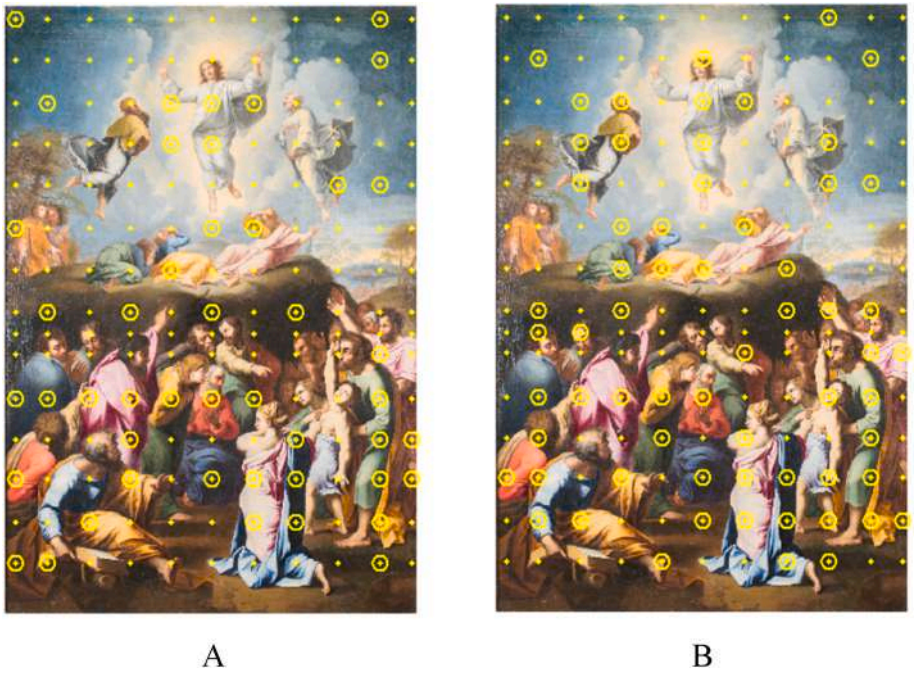

Fig. 2. Locations (circled points) of the preferred scanned points in The Transfiguration selected by Expert #1 (A) and Expert #2 (B).

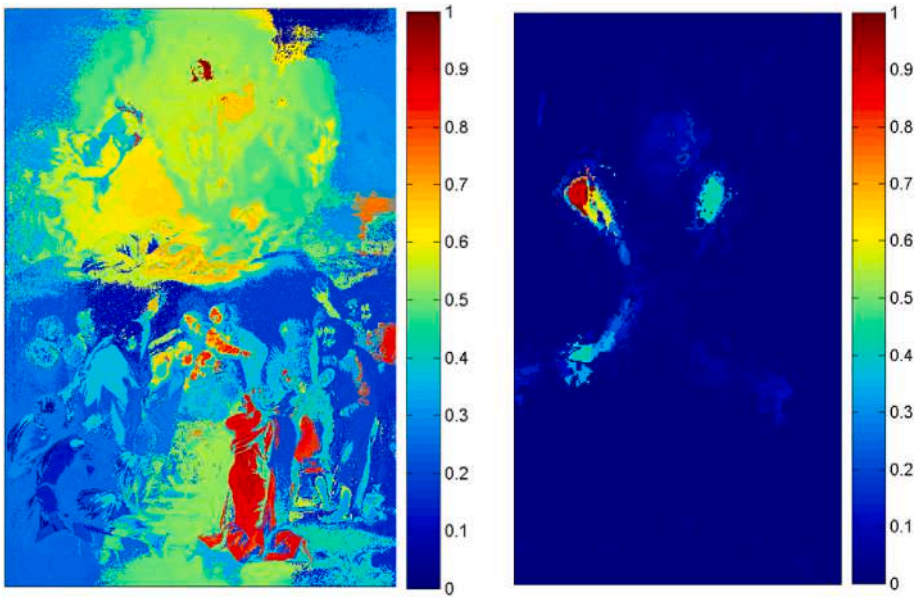

Fig. 3. Chemical element maps generated with the MHD algorithm for Pb on The Transfiguration and Hg on Man. The sizes of the maps are 4456 (height) × 3020 (width) pixels and 357 (height) × 204 (width) pixels, respectively; the same sizes as the RGB images of the paintings.

the quantities of the chemical elements in that point. This means that these approaches cannot be directly applied to regression problems where classes are not available. Nevertheless, generating a chemical element map can be considered as a regression problem [12], i.e. a problem where a continuous response variable is estimated from a set of predictors. In the case of the generation of a chemical element map, the data from the scanned points in the painting would be the predictors, and the amount of each chemical element in every position of the map would be the response variable to estimate.

In the past two decades, machine learning has become a very useful tool for regression problems [13]. Machine learning algorithms solve a regression problem by constructing a model from a set of data, which are the predictors or input variables. Then, this model is used to estimate the value of the numerical response variable or output variable. A large number of machine learning algorithms have been proposed in the literature for solving regression problems [14]. These algorithms can be grouped into families according to the data structure or main principles on which they are based. Some relevant families which we could mention are Linear Regressions, Neural Networks, Support Vector Machines, Regression Trees, Rule-Based Methods, Random Forests and Nearest Neighbours, among many others [15].

The chemical element maps generated from XRF data obtained with portable devices allow the experts to better understand a painting, and their use has increased in recent years. In the present study we aim to improve the accuracy obtained with the MHD algorithm in generating these chemical element maps using machine learning regression

Table 1
Average MSE and R^2 (mean ± standard deviation) achieved on the test data by each algorithm for *The Transfiguration* using experts' information to select the training data in a hold-out validation.

	Expert #1		Expert #2	
Number of training points	46		63	
Number of test points	119		102	
	MSE	R^2	MSE	R^2
MHD	0.037 ± 0.016	0.001 ± 0.004	0.037 ± 0.019	0.007 ± 0.020
Linear Regression	0.028 ± 0.014	0.025 ± 0.044	0.028 ± 0.015	0.032 ± 0.064
Polynomial Regression	0.029 ± 0.014	0.030 ± 0.054	0.026 ± 0.011	0.065 ± 0.112
Random Forest	0.029 ± 0.015	0.049 ± 0.066	0.027 ± 0.014	0.102 ± 0.106
Gradient Boosted Trees	0.033 ± 0.013	0.018 ± 0.037	0.040 ± 0.020	0.043 ± 0.084
M5	0.029 ± 0.016	0.010 ± 0.030	0.028 ± 0.016	0.031 ± 0.061
ML Perceptron	0.041 ± 0.024	0.002 ± 0.007	0.053 ± 0.045	0.000 ± 0.000
SVM	0.033 ± 0.012	0.000 ± 0.000	0.033 ± 0.014	0.000 ± 0.000
kNN	0.026 ± 0.014	0.107 ± 0.113	0.026 ± 0.014	0.071 ± 0.119

algorithms instead. Our proposal is not about how to detect the chemical elements at the scanned points; our goal is to provide better methods than MHD to obtain chemical elements maps from previously scanned points on the paintings, where the identification of the amount of each chemical element at each scanned position is already performed. We are interested not only in improving the performance of the MHD algorithm, but also in providing a methodology based on standard machine learning algorithms which require no special configuration or particularization in order to be applied to different paintings. For this purpose, we first selected a set of standard and well-known machine learning algorithms easily available in common statistical packages. Then we compared the performance of all these algorithms against MHD using two paintings with different features in both the painting itself and the XRF scanning process used. Moreover, the validation of the algorithms was performed in two scenarios: 1) using the data provided by the experts with the selection of scanned points to be used as inputs; and 2) using randomized cross-validation. In this way, we could also assess the effectiveness of the methodology from a statistically solid point of view.

The novelty of our study is twofold: 1) we found methods more accurate than MHD in generating chemical element maps from scattered XRF data of paintings; and 2) we provide solutions based on standard machine learning regression algorithms without requiring special

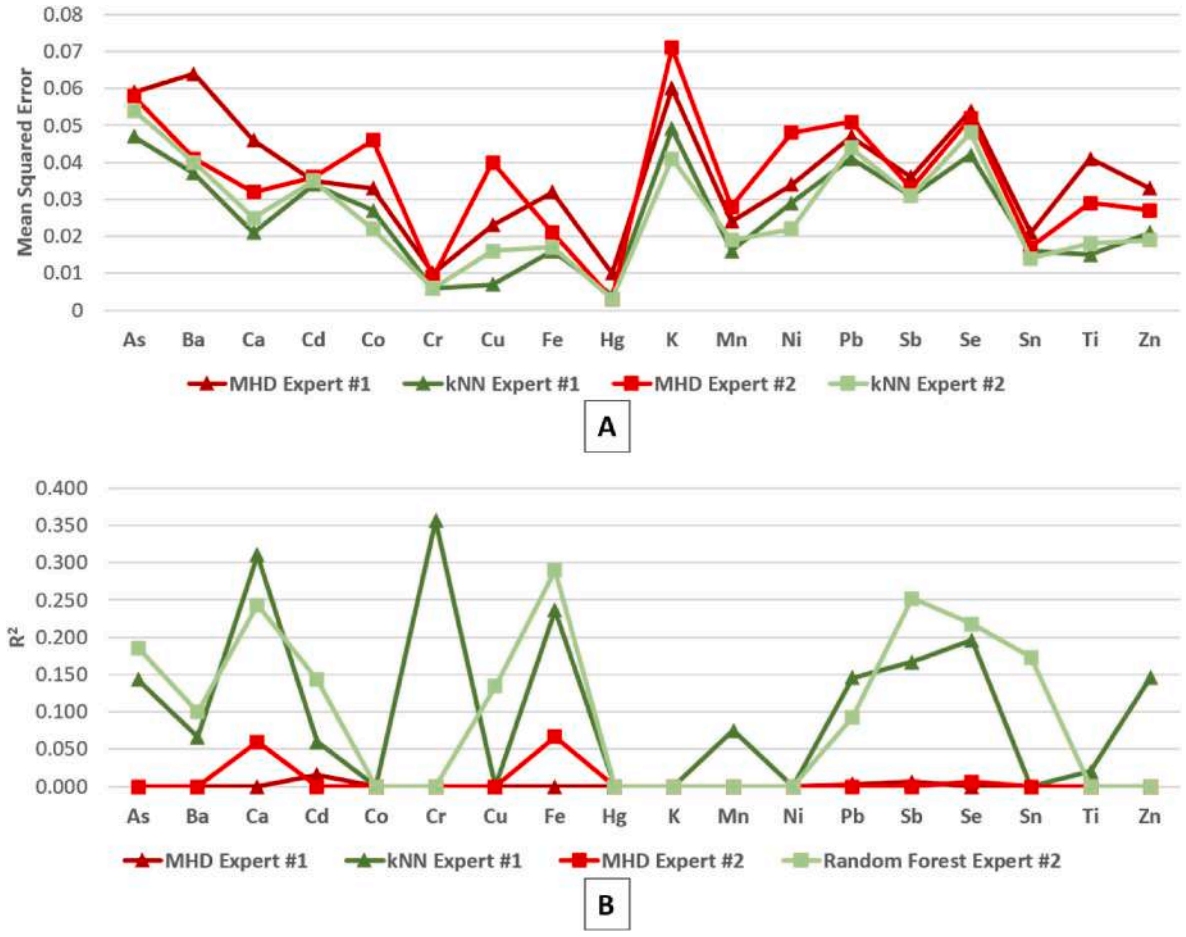

Fig. 4. MSE for MHD and kNN algorithms when performing a hold-out validation for The Transfiguration using as training data the points selected by the experts (A). R^2 for MHD, kNN (expert #1) and Random Forest (expert #2) when performing a hold-out validation for *The Transfiguration* using as training data the points selected by the experts (B).

Table 2
Rankings of the algorithms obtained by the Friedman's test in average MSE and R^2 for *The Transfiguration* using experts' selections as training data in a hold-out validation.

Expert #1				Expert #2			
MSE		R^2		MSE		R^2	
Algorithm	Ranking	Algorithm	Ranking	Algorithm	Ranking	Algorithm	Ranking
kNN	2.81	kNN	7.61	kNN	3.69	Random Forest	6.78
Linear Regression	4.39	Random Forest	6.00	M5	4.28	kNN	6.06
Random Forest	4.86	Polynomial Regression	5.47	Polynomial Regression	4.72	Polynomial Regression	5.81
M5	5.14	Linear Regression	5.22	Random Forest	4.94	Linear Regression	5.11
Polynomial Regression	6.22	Gradient Boosted Trees	4.56	Linear Regression	5.06	Gradient Boosted Trees	5.03
Gradient Boosted Trees	7.69	MHD	4.22	MHD	8.17	M5	4.86
SVM	8.72	M5	4.17	Gradient Boosted Trees	8.22	MHD	4.03
MHD	9.86	ML Perceptron	3.99	SVM	8.69	ML Perceptron	3.67
ML Perceptron	10.72	SVM	3.75	ML Perceptron	10.86	SVM	3.67

Table 3
Adjusted p-values (*Apv*) using Holm's test. MHD versus machine learning algorithms on average MSE and R^2 for *The Transfiguration* using experts' selections as training data in a hold-out validation. NS: no significant.

Expert #1			Expert #2		
Algorithms Comparison	Apv		Algorithms Comparison	Apv	
	MES	R^2		MES	R^2
kNN vs MHD	$p < 0.001$	$p < 0.01$	kNN vs MHD	$p < 0.01$	$p < 0.05$
Linear Regression vs MHD	$p < 0.001$	NS	M5 vs MHD	$p < 0.05$	NS
Random Forest vs MHD	$p < 0.01$	$p < 0.05$	Polynomial Regression vs MHD	$p < 0.05$	$p < 0.05$
M5 vs MHD	$p < 0.01$	NS	Random Forest vs MHD	$p < 0.05$	$p < 0.01$
Polynomial Regression vs MHD	$p < 0.05$	NS	Linear Regression vs MHD	$p < 0.05$	NS
Gradient Boosted Trees	NS	NS	Gradient Boosted Trees	NS	NS
ML Perceptron	NS	NS	ML Perceptron	NS	NS
SVM	NS	NS	SVM	NS	NS

configurations in order to be applied to different paintings. Therefore, our results can be easily reproduced by other researchers in their own studies.

The remainder of the paper is organized as follows: first we describe the paintings used as test data and the different techniques to be compared: MHD and machine learning algorithms. Then we detail the experimental design of our study and the statistical analysis performed to compare the performance of the algorithms. Next we show the results obtained, and finally, we discuss these results and present the conclusions of the study.

2. Material and methods

2.1. Test paintings and data acquisition

Two different paintings were used to test the accuracy of the algorithms: a copy of *The Transfiguration*, by Raphael, and a piece of the scroll *The Miraculous Interventions of Jizō Bosatsu*, by an unknown artist (see Fig. 1A and Fig. 1. C). These two paintings were dated in different periods, elaborated in geographic regions very distanced from each other, and sampled using two different XRF scanners.

The Transfiguration is the last painting by Raphael [16]. The copy of The Transfiguration which we used in this study is an oil painting on canvas belonging to a private collection [5]. In order to process the painting, we used an RGB image of it with dimensions 4456 (height) × 3020 (width) pixels, the canvas size being 93 × 63 cm (see Fig. 1A). The image and the XRF data for a grid of 165 sampled points were provided by the authors of [5]. These data were acquired with a hand-held XRF spectrometer Niton XL3t GOLDD+ (Thermo Fisher Scientific, Waltham, MA, USA) with a silver anode (50 kV, 200 μA) [17]. The analyzer was fitted with a camera and with a suitably equipped Small Spot analyzer, with which the analysis could be restricted to a small area of the camera angle (3 mm). The data for each point consisted of its pixel coordinates and the scanned amounts of the chemical elements As, Ba, Ca, Cd, Co, Cr, Cu, Fe, Hg, K, Mn, Ni, Pb, Sb, Se, Sn, Ti and Zn. The locations of the 165 scanned points are shown in Fig. 1B.

Additionally, Expert #1 (Giacomo Chiari, retired ChiefScientist at the Getty Conservation Institute, Los Angeles, USA) and Expert #2 (author M.R. Blanc) selected two subsets with their preferred points, respectively. These preferred points were the locations of the painting which would have been scanned by the experts in the usual processing methodology, as described next in **Section 2.2**. The locations selected by the experts are shown in Fig. 2.

The Miraculous Interventions of Jizō Bosatsu is a Japanese handscroll by an unknown artist dated to the mid-thirteenth century [18]. This scroll is exhibited in the Freer Gallery of Art (Smithsonian Institution) in Washington D.C. (USA). The dimensions of the scroll are approximately 14 m length and 30 cm width. In this study we used a piece of the whole scroll, what we call *Man*, a scene showing a man wearing garments. The RGB image of *Man* was of dimensions 357 (height) × 204 (width) pixels (see Fig. 1C). The RGB image and the XRF data of the 1314 sampled points were provided as part of the supplementary material of a previous study performed by K.L Rowberg et al. [19]. This data was acquired with a Bruker Tracer 5 g XRF where the Rh anode was set to 40 kV and the anode current at 70 mA. The scanning assay mode was in a continuous cycle collection on an XY-motorized stage [20]. The distance of the device to the surface of the scroll was in a range of 5 ± 3 mm. The data for each point consisted of its pixel coordinates and the scanned amounts of the chemical elements Al, Ar, As, Ca, Cl, Cu, Fe, Hg, K, Mn, Ni, P, Pb, S, Si, Sr, Ti and Zn. The locations of the 1314 scanned points are shown in Fig. 1D. Data from experts with the selection of preferred points is not available for *Man*.

2.2. Creating chemical element maps using the minimum hyperplane distance (MHD) algorithm

Since its release, SmART_Scan [3] has rapidly become a very useful tool with many applications in the cultural heritage field [4–7]. The main reason behind this success is that the MHD algorithm in SmART_Scan allows the users to obtain very reliable chemical element maps of paintings from a reduced number of scanned points acquired with hand-held devices. In this way, the need of using expensive and not easily transportable MacroXRF scanners is avoided.

The data needed by MHD consists of an RGB image of the painting and n scanned points P_i^s. Each scanned point P_i^s is represented by six values: the position of the point in the RGB image, (x_i^s, y_i^s); its RGB color,

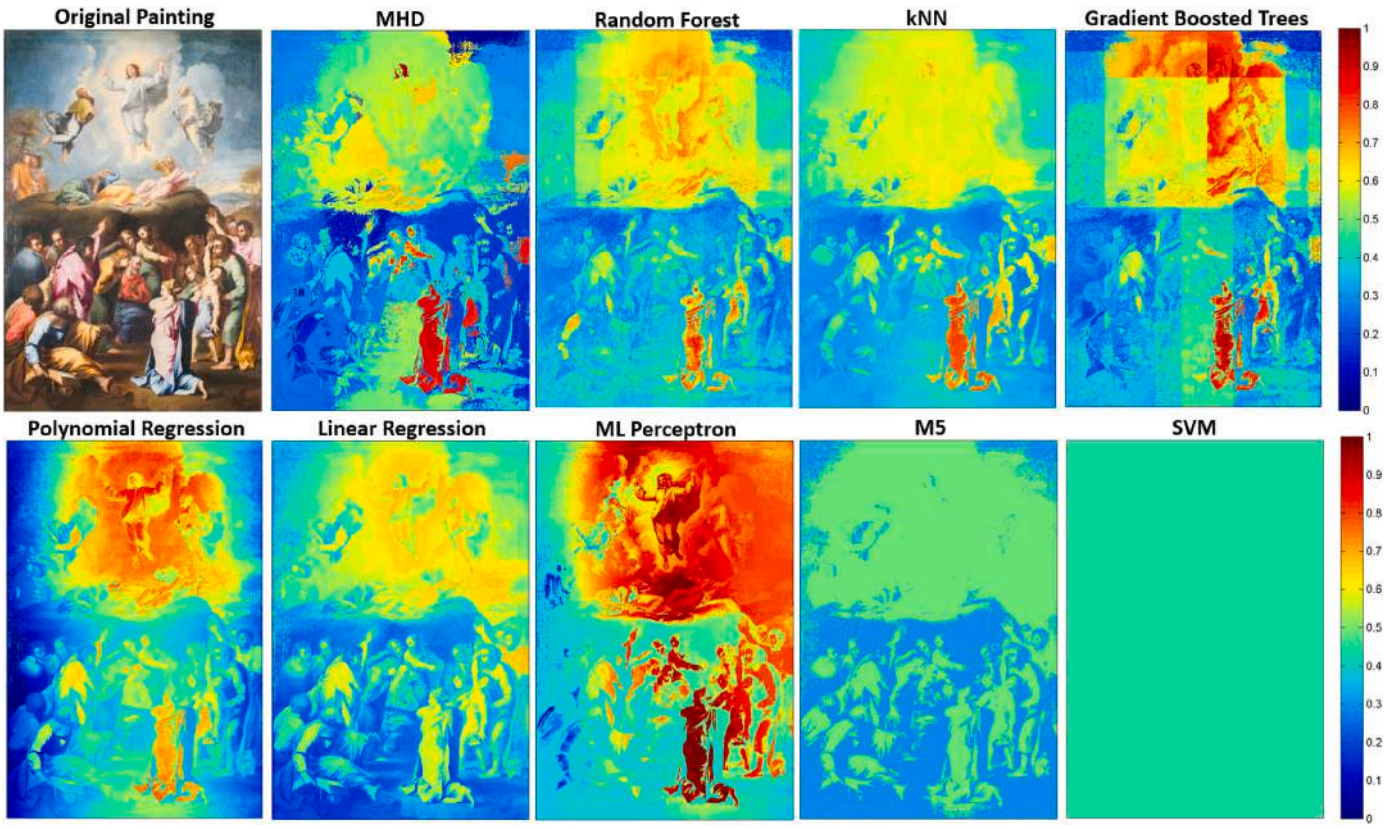

Fig. 5. Chemical element maps for Pb generated with all the algorithms on The Transfiguration using expert #2 data for training (63 points). The size of the maps is 4456 (height) × 3020 (width) pixels, the same size as the RGB image of the painting.

(R_i^s, G_i^s, B_i^s); and the amount of the chemical element scanned in that position, C_i^s. A chemical element map represents the distribution of an element along the painting, so a different map is generated for each chemical element available in the scanning process.

For each point P of the remaining pixels in the RGB image, its position (x, y) and its RGB color (R, G, B) are known. The amount of the chemical element in that point, C, is the only unknown to be calculated. MHD computes C by calculating the Euclidean distance, d_i, between P and the n scanned points P_i^s as follows: $d_i = \sqrt{(x-x_i^s)^2 + (y-y_i^s)^2 + (R-R_i^s)^2 + (G-G_i^s)^2 + (B-B_i^s)^2}$.

And then the value assigned to C at the point P is the amount of chemical element of the scanned point with the minimum of those n computed d_i values. The map is finally generated by repeating this process for every pixel in the RGB image. As an example, Fig. 3 shows the resulting chemical element maps generated with the MHD algorithm for the chemical elements Pb on The Transfiguration and Hg on Man.

2.3. Machine learning regression algorithms

Many machine learning algorithms have been proposed for solving regression problems [14,15]. Usually these algorithms focus on specific areas or problems, with custom processing and configurations. One goal of our study is to provide a methodology based on standard and easily accessible algorithms. This is the reason why our approach is based on a selection of eight algorithms chosen among the most well-known machine learning techniques which are available in standard statistical tools. We also selected the algorithms belonging to different classical families of machine learning techniques in order to cover a wide range of methodologies. Based on these criteria, the algorithms used in our study were Linear Regression [21], Polynomial Regression [22], Random Forest [23], Gradient Boosted Trees [24], M5 [25], Multi Layer (ML) Perceptron [26], Support Vector Machine (SVN) [27] and k-nearest neighbour (kNN) [28]. All these algorithms are standard and very well known among the scientific community. Nevertheless, the reader who needs more details on these techniques can consult the previously-referred papers for each algorithm or general references in data mining such as [15].

The statistical software tool used in our study was KNIME [29]. We selected KNIME because it is a visual software tool which allows the user to perform machine learning analysis easily by creating the processing pipeline visually through predefined modules. Next, we briefly describe the parameters configuration in KNIME for each algorithm.

- *Linear Regression.* No parameter configuration is available in KNIME. Simple linear regression using least squares was performed.
- *Polynomial Regression.* The algorithm was configured in KNIME to perform polynomial regression of order two.
- *Random Forest.* The algorithm was configured in KNIME with the default parameters provided by the tool: no limit for tree depth, no minimum node size and generation of 100 models. Five different seeds were used to avoid biases due to the random process in the generation of the models.
- *Gradient Boosted Trees.* The default configuration in KNIME was used: tree depth limited to four levels, learning rate of 0.1 and generation of 100 models.
- *M5.* The M5 algorithm was configured with the default values provided by KNIME: minimum number of instances set to 4, pruning tree and smoothing predictions.
- *ML Perceptron.* The default parameters used were: 0.3 for learning rate, 0.2 for momentum, training time of 500 epochs, 0% of validation set size, validation threshold of 20 times, and a value of number of attributes/2 for the number of hidden layers.
- *SVM.* The default parameters in KNIME for the SVM regression algorithm were: radial basis kernel with a gamma value of 0.0, 1.0 for

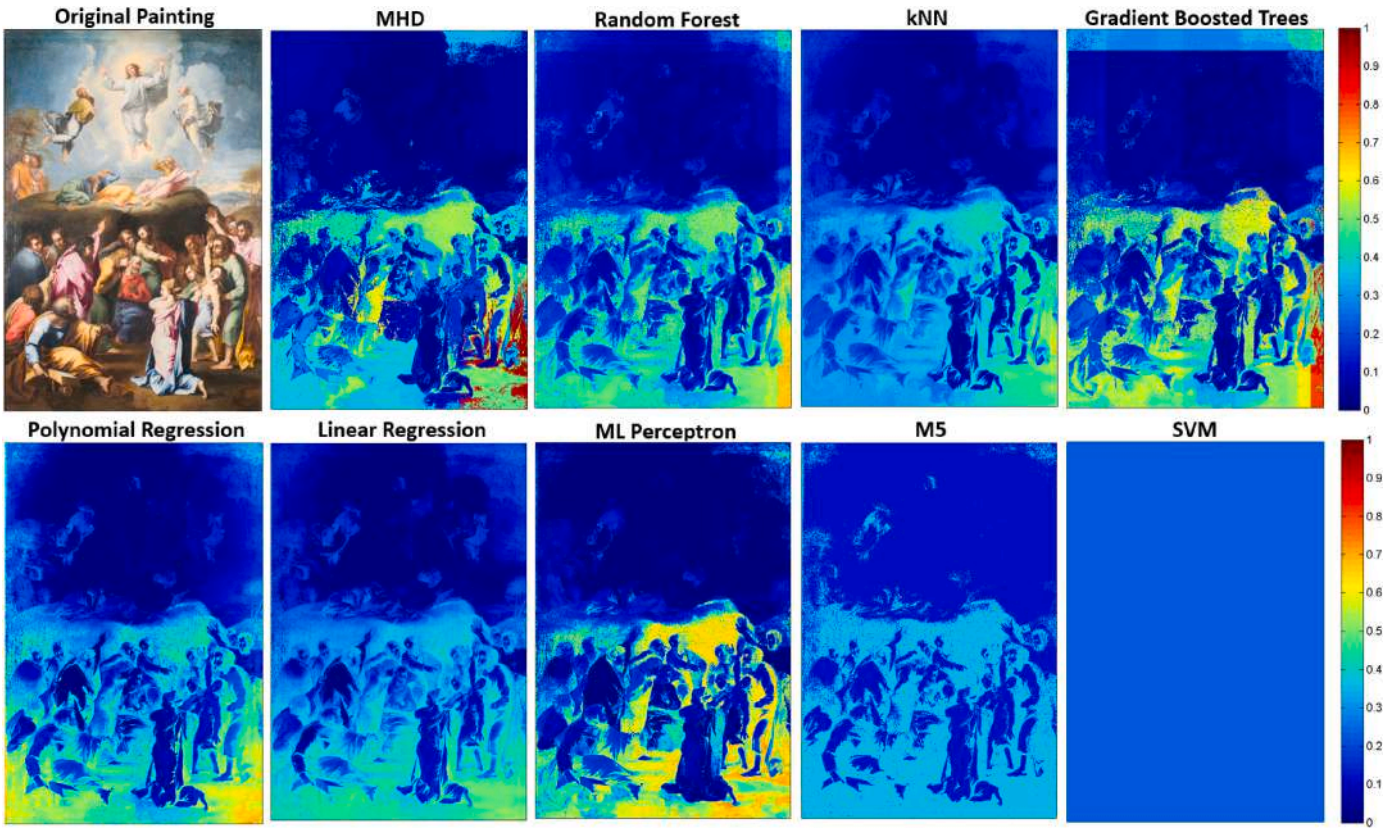

Fig. 6. Chemical element maps for Ca generated with all the algorithms on The Transfiguration using expert #2 data for training (63 points). The size of the maps is 4456 (height) × 3020 (width) pixels, the same size as the RGB image of the painting.

Table 4
Average MSE and R^2 (mean ± standard deviation) achieved in the test data for each algorithm on *The Transfiguration* using a 5-fold cross-validation.

Number of training points	132 *5	
Number of test points	33 * 5	
	MSE	R^2
MHD	0.034 ± 0.011	0.015 ± 0.035
Linear Regression	0.026 ± 0.011	0.117 ± 0.122
Polynomial Regression	0.023 ± 0.007	0.170 ± 0.166
Random Forest	0.021 ± 0.007	0.234 ± 0.195
sGradient Boosted Trees	0.024 ± 0.008	0.190 ± 0.176
M5	0.027 ± 0.010	0.111 ± 0.128
ML Perceptron	0.035 ± 0.017	0.034 ± 0.068
SVM	0.034 ± 0.012	0.000 ± 0.000
kNN	0.024 ± 0.008	0.159 ± 0.156

the cost parameter, an epsilon value of 0.001 for the tolerance of the termination criterion, and an epsilon value of 0.1 for the loss function.

- *kNN*. We used the default parameter in KNIME for the search algorithm, a linear search based on the Euclidean distance. The value of *k* was manually set to 10 because the default value provided by KNIME is 1, which is a non-useful value in our case for performing regression.

Our goal is to obtain a methodology based on automatic processing without expert intervention and regardless of the painting, so we used the default parameters of the algorithms in KNIME because adjusting them is not a trivial task and would depend on the painting being studied. We know that algorithms such as ML Perceptron or SVM would obtain better results if a fine selection of the parameters were performed carefully. But, as mentioned before, this would imply the need of additional experts in machine learning and to perform changes in the configurations depending on the painting.

2.4. Experimental design and statistical analysis

We compared MHD versus machine learning algorithms in three different experiment configurations.

1) *The Transfiguration* using information from the experts and a hold-out validation: In this experiment we tested the capability of machine learning algorithms to obtain accurate maps by using as training data the same information provided by experts in the MHD algorithm. The data provided by the experts are very reduced in number, a non-desirable situation when training machine learning models. The validation was performed by using hold-out [30] because it reproduces the way in which the data selected by the experts is used to obtain the map. The selected points are used to interpolate the remainder points of the map in MHD. In our hold-out validation, the training data are the expert-selected points, and the test data are the remainder of the scanned points.

Note that MHD is not a machine learning technique but an interpolation algorithm, so the terms "train data" and "predicting" make no sense for this algorithm; "inputs" and "interpolating", respectively should be used instead. In any case, for the sake of simplicity, we will use these terms for both machine learning algorithms and MHD equally throughout the text.

2) *The Transfiguration* without using experts' information and validating through cross-validation: With this experiment we wanted to demonstrate the usefulness of our methodology using well-known

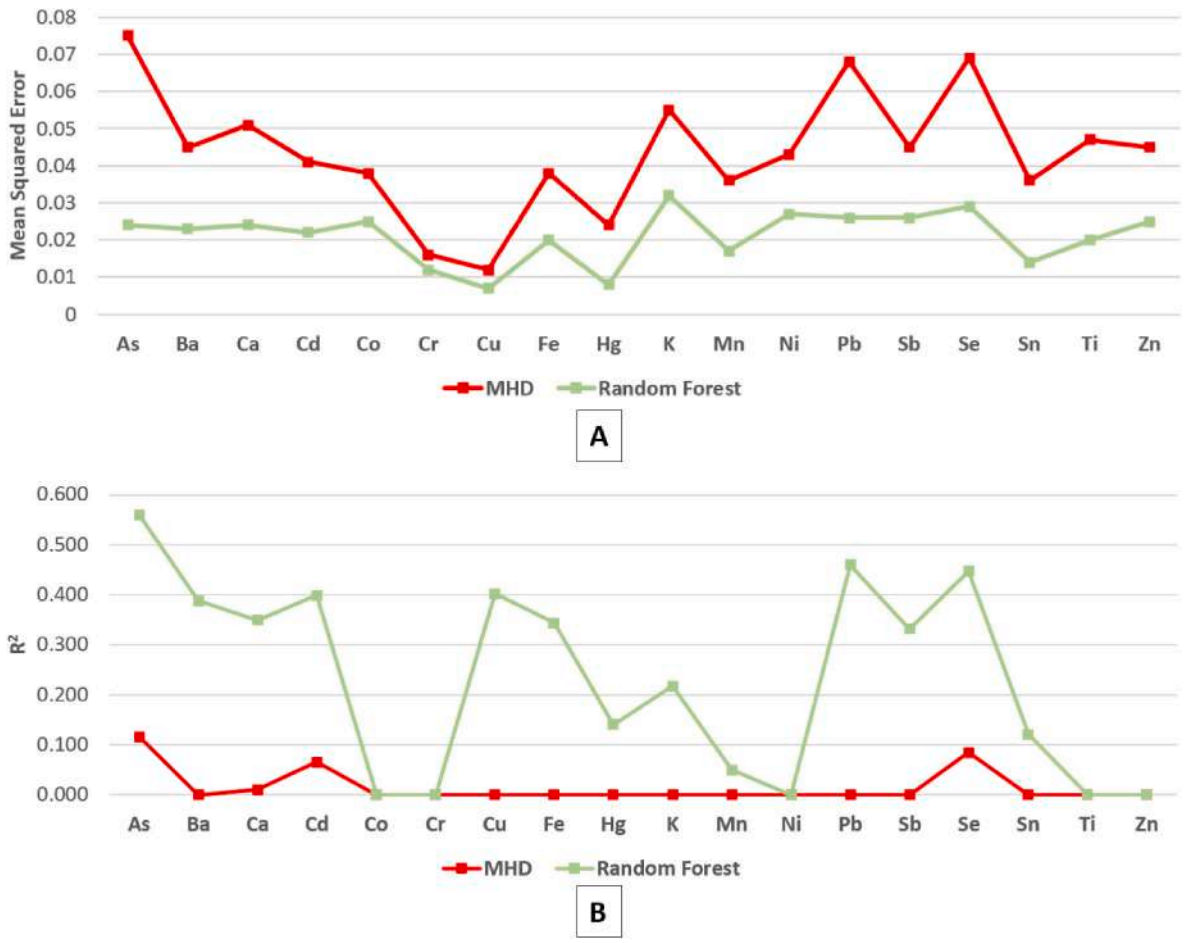

Fig. 7. MSE (A) and R^2 (B) for MHD and Random Forest algorithms when performing a 5-fold cross-validation for *The Transfiguration*.

Table 5
Rankings of the algorithms obtained by Friedman's test on average MSE and R^2 in the test data for *The Transfiguration* using a 5-fold cross-validation.

MSE		R^2	
Algorithm	Ranking	Algorithm	Ranking
Random Forest	2.166	Random Forest	7.472
Polynomial Regression	3.750	Polynomial Regression	6.888
Gradient Boosted Trees	4.666	Gradient Boosted Trees	6.055
kNN	4.805	kNN	5.805
Linear Regression	4.888	Linear Regression	5.583
M5	5.944	M5	4.416
MHD	9.666	ML Perceptron	3.166
SVM	10.083	MHD	2.972
ML Perceptron	10.333	SVM	2.638

Table 6
Adjusted p-values (*Apv*) using Holm's test. MHD versus machine learning algorithms on average MSE and R^2 for *The Transfiguration* using a 5-fold cross-validation. NS: no significant.

Algorithms Comparison	Apv	
	MSE	R^2
Random Forest vs MHD	$p < 0.001$	$p < 0.001$
Polynomial Regression vs MHD	$p < 0.001$	$p < 0.001$
Gradient Boosted Trees vs MHD	$p < 0.01$	$p < 0.05$
kNN vs MHD	$p < 0.01$	$p < 0.05$
Linear Regression vs MHD	$p < 0.01$	NS
M5 vs MHD	$p < 0.05$	NS
ML Perceptron vs MHD	NS	NS
SVM vs MHD	NS	NS

statistical techniques used to estimate how accurately a predictive model will perform in practice. In this way MHD and machine learning algorithms were validated through a 5-fold cross-validation [31], where exactly the same randomly-generated folds were used for testing all the algorithms, including MHD. We used 5-fold cross-validation, instead of the more typical 10-fold cross-validation used in classification, trying to avoid the problem of overfitting in regression [14,32]. For the case of Random Forest, we tested the algorithm five times using different seeds in the generation of each tree in order to avoid biases due to a specific seed selection. The final results for Random Forest were the mean of these five executions.

3) *Man* with cross-validation: Once we studied the effectiveness of the proposed methodology in a painting for which experts' information is available, the next step was to test how the methodology worked in another painting with two differentiating features regarding *The Transfiguration*: 1) a much larger dataset of scanned points, and 2) no experts' information available for deciding whether certain scanned points are relevant or not. Following the same reasoning as previously explained, the accuracy of all the methods was performed using a 5-fold cross-validation.

The goal of these three experiments was to prove the effectiveness of the methodology not only in paintings where expert input is available, but also in paintings where the only information provided is the pigment at several sampled locations. To the best of our knowledge, this is a new solution for the problem of generating chemical element maps of paintings.

Table 7
Average MSE and R^2 (mean ± standard deviation) achieved in the test data for each algorithm on *Man* using a 5-fold cross-validation.

Number of train points	1051 *5	
Number of test points	263 * 5	
	MSE	R^2
MHD	0.022 ± 0.020	0.251 ± 0.324
Linear Regression	0.017 ± 0.011	0.199 ± 0.208
Polynomial Regression	0.016 ± 0.010	0.239 ± 0.229
Random Forest	0.013 ± 0.011	0.388 ± 0.314
Gradient Boosted Trees	0.013 ± 0.011	0.351 ± 0.322
M5	0.014 ± 0.009	0.310 ± 0.270
ML Perceptron	0.018 ± 0.012	0.183 ± 0.259
SVM	0.025 ± 0.013	0.000 ± 0.000
kNN	0.013 ± 0.010	0.353 ± 0.300

Two different and complementary metrics have been used in order to measure the accuracy of the algorithms: mean squared error (MSE) and coefficient of determination (R^2). MSE is calculated as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(pred\_C(x_i, y_i, R_i, G_i, B_i) - C_i)^2$$

where N is the number of test points, $pred\_C(x_i, y_i, R_i, G_i, B_i)$ is the prediction of the chemical element obtained from the model generated by the algorithm for the test point i, and C_i is the known value of the chemical element for the point i.

Since the value of each chemical element in each scanned point was normalized between 0 and 1, then the MSE values obtained were also between 0 and 1, which makes these results easily interpretable.

On the other hand, the coefficient of determination, R^2, is defined as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Table 8
Rankings of the algorithms obtained by the Friedman test on *Man* using a 5-fold cross-validation.

MSE		R^2	
Algorithm	Ranking	Algorithm	Ranking
Random Forest	2.722	Random Forest	7.333
kNN	3.638	Gradient Boosted Trees	6.583
M5	3.638	kNN	6.416
Gradient Boosted Trees	3.777	M5	6.333
Polynomial Regression	4.361	Polynomial Regression	5.333
Linear Regression	5.444	Linear Regression	4.416
ML Perceptron	6.805	MHD	3.805
MHD	6.888	ML Perceptron	3.000
SVM	7.722	SVM	1.777

Table 9
Adjusted p-values (*Apv*) using Holm's test. MHD versus machine learning algorithms on MSE for *Man* using a 5-fold cross-validation. NS: no significant.

Algorithms Comparison	Apv	
	MSE	R^2
Random Forest vs MHD	$p < 0.001$	$p < 0.001$
kNN vs MHD	$p < 0.001$	$p < 0.01$
M5 vs MHD	$p < 0.001$	$p < 0.05$
Gradient Boosted Trees vs MHD	$p < 0.01$	$p < 0.01$
Polynomial Regression vs MHD	$p < 0.05$	NS
Linear Regression vs MHD	NS	NS
ML Perceptron vs MHD	NS	NS
SVM vs MHD	NS	NS

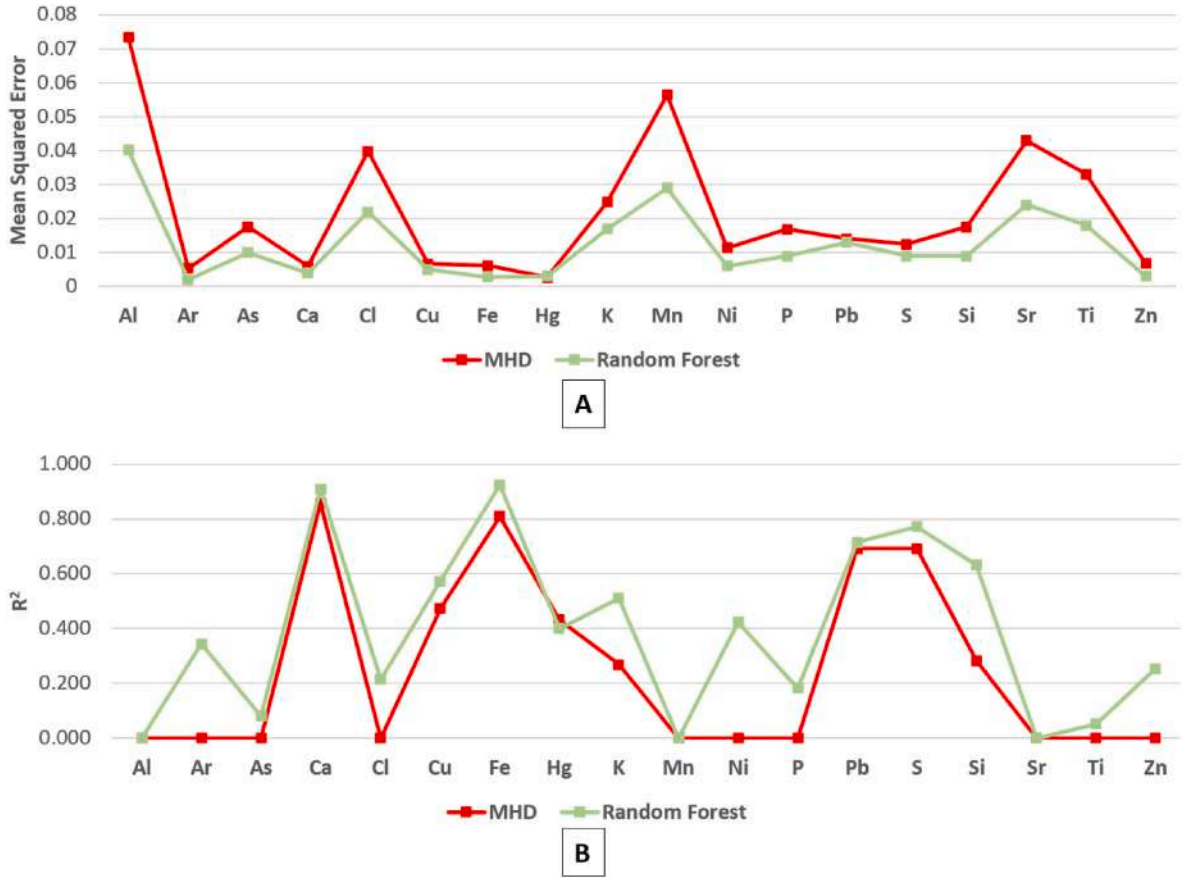

Fig. 8. MSE (A) and R^2 (B) for MHD and Random Forest algorithms when performing a 5-fold cross-validation for *Man*.

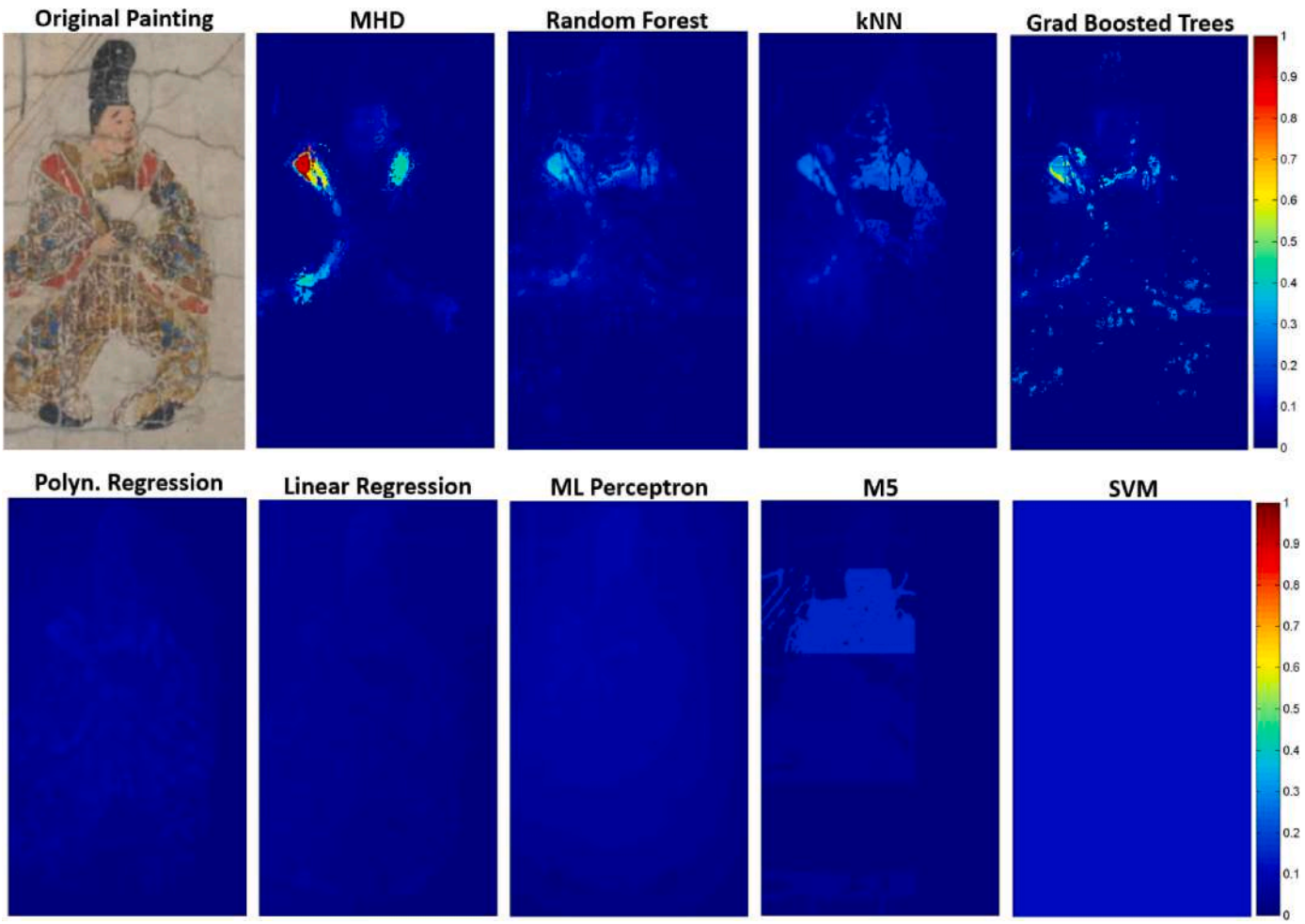

Fig. 9. Chemical element maps for Hg generated with all the algorithms on Man using as training data the points of the first fold in cross-validation (1051 points). The size of the maps is 357 (height) × 204 (width) pixels, the same size as the RGB image of the painting.

where $SS_{res} = \sum_{i=1}^{N} (pred\_C(x_i, y_i, R_i, G_i, B_i), -C_i)^2$ and $SS_{tot} = \sum_{i=1}^{N} (C_i - \overline{C})^2$, being $\overline{C}$ the mean of the values C_i. R^2 measures the proportion of the variation in the prediction (the amount of chemical element) that can be explained from the predictors (position and color). R^2 ranges between 0 and 1, where 1 means a perfect prediction.

Statistically significant differences in average MSE and R^2 between algorithms were computed by using, jointly, the non-parametric Friedman's and Holm's tests [33,34]. First, we computed the ranking of each algorithm using the Friedman's test [35], and then a post-hoc Holm's test [36] was applied to obtain the adjusted *p*-values counting for multiple pairwise algorithm comparisons.

3. Results

In the next three sections we show the results of the analysis performed for each proposed experiment.

3.1. Results for The Transfiguration using experts' information

Table 1 shows the average MSE and R^2 values obtained for each algorithm when using a hold-out validation where the training data are the points selected by the experts. These average MSE and R^2 values correspond to the mean of the MSE and R^2 values obtained by predicting the amount of each one of the 18 chemical elements present in the remaining scanned points of the painting.

Fig. 4 shows the graphical comparison of individual MSE and R^2 obtained for MHD versus the best machine learning algorithm in each experiment (kNN for Expert #1 and Random Forest for Expert #2) for all chemical elements in the painting.

Table 2 shows the rankings using Friedman's test in average MSE and R^2 for all the algorithms in both experiments. The kNN algorithm obtained the best ranking except for R^2 when trained by expert #2, where Random Forest is the best ranked algorithm. Several machine learning algorithms outperformed the MHD algorithm in both experiments.

Finally, Table 3 shows the adjusted p-values obtained using Holm's test when comparing the machine learning methods versus MHD in average MSE and R^2. Significant differences, compared to MHD, were found in MSE for kNN, Linear Regression, Random Forest, M5 and Polynomial Regression in both experiments. In the case of R2, significant differences were found for kNN (experts #1 and #2), Random Forest (experts #1 and #2) and Polynomial Regression (only expert #2).

These results show that kNN and Random Forest significantly outperform MHD in both MES and R^2 when trained with data selected by experts. Figs. 5 and 6 show examples of the generated maps with all the algorithms for two of the main chemical elements present on The Transfiguration, Pb and Ca, respectively. Pb (white lead) is present in white and light colors, such as clouds and light clothes in the painting. On the other hand, Ca (bone black) is present in dark zones, such as the background without characters of the lower half of the painting. These maps were generated using as training data the points selected by expert #2 (63 points). A visual inspection of the obtained maps confirms that R^2 is a reliable metric for discarding algorithms that, even obtaining low average MSE scores, produce useless maps, such as the case of SVM (R^2

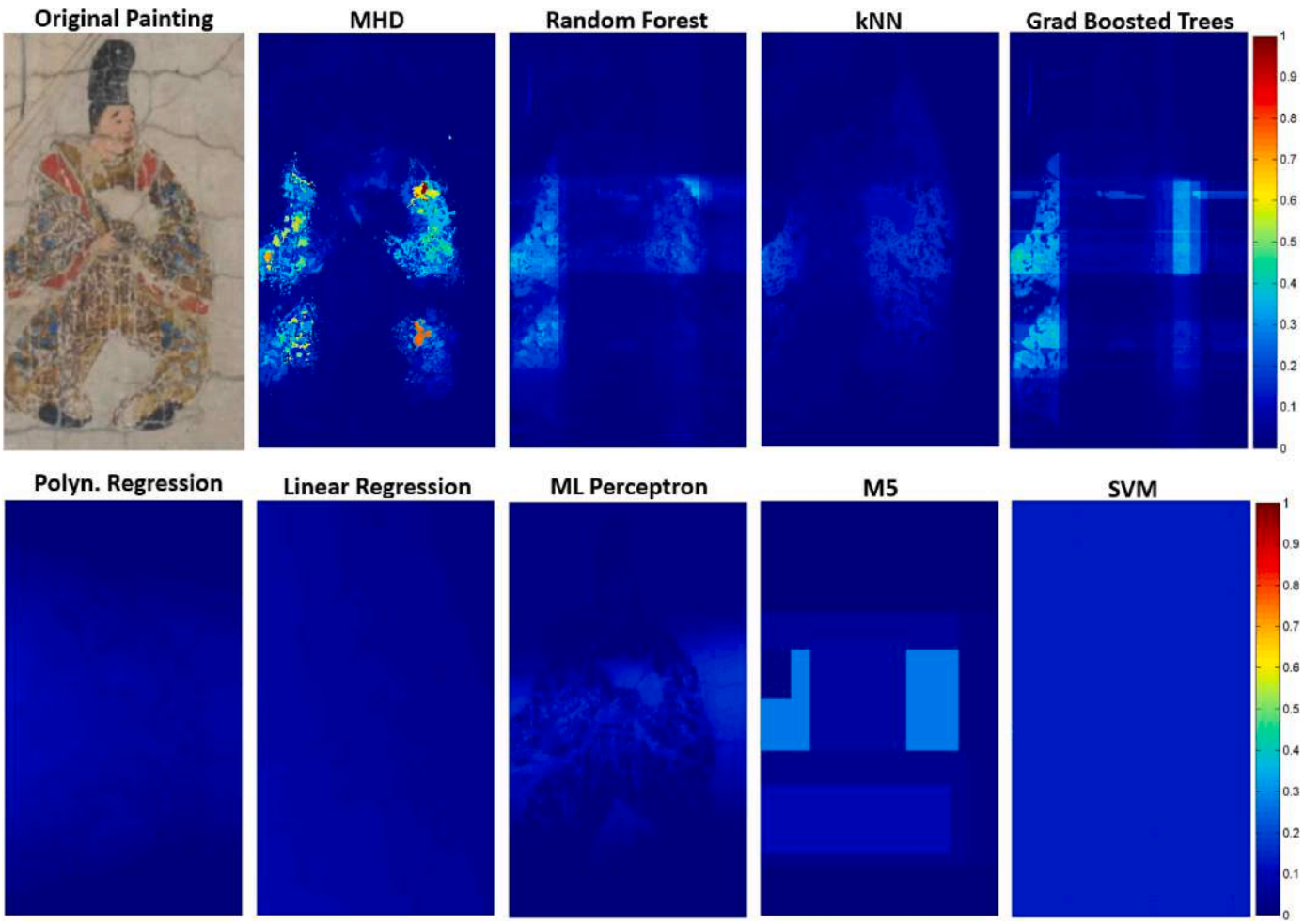

Fig. 10. Chemical element maps for Cu generated with all the algorithms on Man using as training data the points of the first fold in cross-validation (1051 points). The size of the maps is 357 (height) × 204 (width) pixels, the same size as the RGB image of the painting.

Table 10
Times, in seconds, for training and predicting a chemical element map (Zn). Training times are the average time obtained in the corresponding 5-fold cross-validation.

	The Transfiguration		Man	
Training points	132		1051	
Prediction points	13,457,120		72,828	
	Training (s)	Prediction (s)	Training (s)	Prediction (s)
MHD	–	19.853	–	0.988
Linear Regression	0.039	22,591.649	0.097	11.104
Polynomial Regression	0.020	18,105.943	0.033	5.925
Random Forest	0.215	30,625.142	0.760	78.366
Gradient Boosted Trees	0.181	14,068.807	0.397	17.085
M5	0.141	10,072.450	0.368	6.341
ML Perceptron	0.230	9256.867	0.878	3.987
SVM	0.074	13,742.574	0.297	25.863
kNN	0.023	12,682.316	0.028	29.186

$= 0.000 \pm 0.000$).

3.2. Results for The Transfiguration using cross-validation

Table 4 shows the average MSE and R^2 values obtained for each algorithm when using a 5-fold cross-validation, i.e. we randomly split the dataset of scanned points into five folds, each containing 20% of the points of the dataset, where four folds were used for training and one for testing. Note that results for Random Forest are shown as the average value for five executions with different seeds.

Fig. 7 shows the graphical comparison of individual MSE and R^2 obtained for MHD versus the best machine learning algorithm, Random Forest, for all chemical elements in *The Transfiguration*.

Table 5 shows the rankings using Friedman's test in average MSE and R^2 for all the algorithms. The Random Forest algorithm obtained the best ranking, and again several machine learning algorithms outperformed the MHD algorithm.

Table 6 shows the adjusted p-values obtained using Holm's test when comparing the machine learning methods versus MHD in average MSE and R^2.

These results show that Random Forest, Polynomial Regression, Gradient Boosted Trees and kNN significantly outperform MHD in both MSE and R^2 when a cross-validation was used.

3.3. Results for man

Table 7 shows the average MSE and R^2 values obtained for each algorithm when using a 5-fold cross-validation in predicting the amounts of the chemical elements in *Man*. Results for Random Forest are the average value for five executions with different seeds.

Fig. 8 shows the graphical comparison of individual MSE and R^2 obtained for MHD versus the best machine learning algorithm, Random Forest, for all chemical elements in *Man*.

Table 8 shows the rankings using Friedman's test in average MSE and R^2 for all the algorithms. The Random Forest algorithm obtained the best ranking, as in the case of The Transfiguration with 5-fold cross-validation. Also, all machine learning algorithms except SVM and ML Perceptron outperformed the MHD algorithm.

The adjusted p-values obtained using Holm's test when comparing the machine learning methods versus MHD in average MSE and R^2 are shown in Table 9.

These results show that Random Forest, M5, kNN and Gradient Boosted Trees outperformed MHD in both MSE and R^2 when predicting the amounts of the chemical elements in *Man* using cross-validation. Figs. 9 and 10 show examples of the generated maps with all the algorithms for two of the key chemical elements present on Man, Hg and Cu, respectively. Hg (Vermilion or Cinnabar) is present in the red orange zones of the dress in the painting, while Cu (Azurite and/or Malachite) is present in the blue areas of the dress. These maps were generated using as training data the points included in the first fold used in the cross-validation (1051 points). As in the case of The Transfiguration (see Figs. 5 and 6), algorithms with very low R^2, such as SVM ($R^2 = 0.000 \pm 0.000$), produce maps which provide no useful information.

3.4. Performance analysis

We also assessed the performance of all algorithms regarding the execution time of both training and predicting (generating) a chemical element map. Table 10 shows the timing for both paintings, *The Transfiguration* and *Man*. Training times correspond to the average time obtained for the 5 training processes in the 5-fold cross-validation. Prediction times are the total times in generating a chemical element map (Zn) from the corresponding trained model, that is, predicting the quantity of Zn in each point (pixel) of the painting. These times were obtained in a laptop equipped with a CPU Intel Celeron N4020 at 2.8 GHz with two cores, accompanied by 8 GB RAM memory and a GPU Intel UHD 600 with 3 GB of RAM memory and 12 execution units. MHD is an interpolation algorithm, so no training time is needed because the training points are used directly as input in the interpolation process.

Results shows that MHD is much faster than machine learning algorithms when large chemical element maps are generated (around 20 s for MHD vs. 2.5–8.5 h for machine learning algorithms for the case of *The Transfiguration*). This high difference in processing time has a simple explanation: the implementation of the MHD algorithm that we used was optimized for using GPU acceleration, while the implementation of the machine learning algorithms in KNIME did not allow us to use hardware acceleration. However, other software packages, such as cuML [37], implement machine learning algorithms to run on GPU. So, using this kind of optimized machine learning algorithms could greatly improve the performance in a final production scenario.

Nevertheless, the most relevant result is the very fast training times achieved for most machine learning algorithms. All training times are below 1 s, which are insignificant compared to prediction times. This result is of great relevance because it means that the retraining process needed when analyzing a new painting is completely inexpensive compared to the time for generating the chemical element maps.

4. Discussion

In this study we have assessed the ability of standard machine learning regression algorithms to improve the accuracy in generating chemical element maps of paintings compared to the MHD algorithm. This algorithm is used by experts who must select a set of scanned points to interpolate the amounts of chemical elements in the remaining points of the map. We wanted to remove the necessity of expert intervention for setting the machine learning algorithms, so we chose well-known algorithms with their default configuration in widely used statistical tools such as KNIME. Standard statistical techniques such as hold-out and cross-validation were used to compare the accuracy of the algorithms in two complementary metrics: MSE and R^2. We used non-parametric tests for the statistical comparisons of the algorithms: Friedman's test for ranking the algorithms and a post-hoc Holm's test to find the pairwise comparisons producing significant differences. Below we discuss the results obtained for each experiment.

4.1. MHD versus machine learning algorithms when experts' information is provided and used as training data

Results in **Section 3.1.** showed that kNN and Random Forest performed significantly better than MHD when the accuracy of the predictions, in terms of MSE and R^2, was tested using a hold-out validation in which the training data was the points selected by the two experts. These results are remarkable because a very limited number of points were used as training data (46 and 63, respectively), such a reduced number of points being a typical limiting factor in machine learning algorithms.

The lowest average MSE values for both experiments were obtained by the kNN algorithm, a classical algorithm which searches for minimum distances among the neighbours. Recent surveys on machine learning algorithms for regression [14] showed that, in general terms, the best methods would be Random Forest or M5. Nevertheless, kNN outperformed those algorithms in MSE mainly because the training data were very limited in number. R^2 analysis showed that Random Forest achieved the best score when training data was provided by expert #2. This result is in line with the previous hypothesis since the number of points selected by expert #2 is greater than the number for expert #1.

4.2. MHD versus machine learning algorithms with cross-validation: The Transfiguration

Results in **Section 3.2.** showed that Random Forest, Polynomial Regression, Gradient Boosted Trees and kNN performed significantly better than MHD in MSE and R^2 when the accuracy of the predictions was tested using a 5-fold cross-validation. These results indicate that those machine learning algorithms were able to achieve lower errors and higher R^2 values than MHD when tested using randomized positions as training data. This is a key goal in our proposal, since we want to provide methods which perform well from a solid statistical point of view.

The lowest average MSE value and greater R^2 score were achieved by the Random Forest algorithm, which is a result in line with recent rankings of machine learning algorithms for regression [14]. Random Forest also achieved very low deviations in the MSE values, a very desirable feature proving the good performance of the algorithm for almost all chemical elements.

R^2 scores for randomized cross-validation were better than those obtained when the machine learning algorithms were trained with data provided by the experts in a hold-out validation (see Tables 1 and 5, and Figs. 4 and 7). The reason is that training data was much larger when cross-validation was used (132 points) than in the experiment with expert's data (46 points for expert #1, and 63 points for expert #2). Therefore, even when a randomized selection of training points was performed with cross-validation, as the number of training points is high, these points covered the positions with the most relevant

information about colors and chemical elements of the painting. However, MHD algorithm did not improve the R^2 score so clearly when the number of training points increased when using cross-validation. The reason behind this situation is that the values of the chemical element estimated by MHD for generating the map (values at test points) are values of the chemical element contained in the positions of the training points (see Section 2.2 and generated maps with MHD in Figs. 5 and 6). Therefore, for obtaining high R^2 scores, MHD needs a number of scanned positions which covers well the range of possible values of the chemical elements present on the painting.

4.3. MHD versus machine learning algorithms with cross-validation: man

Results in **Section 3.3.** have proved that four machine learning algorithms, Random Forest M5, kNN and Gradient Boosted Trees, also significantly improved the accuracy in predicting the amount of chemical elements for each point of *Man* compared to MHD. This fact confirms the good results obtained for *The Transfiguration*, but now in a very different scenario where many more scanned points are available for training and testing, selected in a randomized way through cross-validation. As a result, we can conclude that those machine learning algorithms allowed us to obtain significantly more accurate chemical element maps than MHD.

As shown in Fig. 8, R^2 scores for MHD and Random Forest on *Man*, although statistically different in average, are much closer than in the case of *The Transfiguration* (see Fig. 7). Cross-validation in *Man* used training sets of 1051 points. This much larger amount of training points together with a less range of different colors in *Man* compared to *The Transfiguration* caused that MHD performed better, although still worse than Random Forest.

Random Forest has clearly achieved the best results for both *The Transfiguration* and *Man* when tested using cross-validation. Therefore, Random Forest would be positioned, in general, as the preferred algorithm for generating the chemical element maps. In case a low number of scanned positions provided by the experts were available, doubts regarding which algorithm should be used could arise, since kNN obtained the best results for *The Transfiguration* in both experiments in which the training data was selected by experts. Nevertheless, the Holm's tests comparing kNN versus Random Forest on average MSE and R^2 provided non-significant p-values in both experiments, so no significant differences in accuracy were found between both algorithms.

4.4. Qualitative assessment of generated maps

Figs. 5 and 6 showed the generated maps with all the algorithms for two main chemical elements present in The Transfiguration, Pb and Ca, respectively. According to a visual assessment of the maps, and in concordance with the MSE and R^2 scores obtained by the algorithms, kNN and Random Forest showed better results than MHD. Pb (white lead) is present in white and light colors, such as clouds and light clothes in the painting. The map generated with MHD (see Fig. 5) showed no presence of Pb in some light zones, as in the case of the arm of the character in the lower-left area of the painting. Also, some dark zones of the painting, such as the lower-central area, should present less amount of Pb than that showed in the map generated with MHD. Since the MHD algorithm assigns to generated points only values contained in the training data, several zones have maximum values of Pb (red zoned in the map) which appears to be not correct: Jesus Christ's hair, boy's skirt, the central character's sleeve and above the little tree on the right edge of the painting. Ca (bone black) is present in dark zones, such as the background without characters of the lower half of the painting. In this case (see Fig. 6), Random Forest map appears to represent better than MHD those dark zones of the lower-left zones of the painting.

Nevertheless, visual inspection of generated maps for Man shown in Figs. 9 and 10 reveals that MHD produced better results than Random Forest, kNN. Hg (orange red areas with vermilion) and Cu (blue zones with azurite) are more clearly delimited in the maps generated with MHD. In fact, the only case in which the R^2 score is better for MHD than Random Forest occurs when generating the map for Hg (see Fig. 8). These good results of MHD in both R^2 scores and visual assessment are caused by the high number of training points used to generate the maps (1051 scanned points), which increases widely the range of values available by the algorithm for assigning the new positions of the map. Unfortunately, counting with that large amount of scanned XRF positions is very difficult and not common when using portable devices for analysing paintings. Therefore, when a limited number of scanned positions are available, our results suggest that the machine learning algorithms Random Forest and kNN are the best choices.

5. Conclusions

Our results suggest that several machine learning regression algorithms significantly improved the MHD algorithm. Random Forest and kNN performed significantly better than MHD when the training data were the points selected by the experts. Moreover, Random Forest and kNN were able to learn and achieve significantly lower average errors and greater R^2 values than MHD when randomized cross-validations were performed in both paintings. The execution time of the training process is insignificant compared to the time for generating the chemical element maps, what allow us to retrain very fast the machine-learning models for a new painting, avoiding the use of complex deep-learning models based on large training datasets of paintings. We can conclude that, in general terms and specially when a limited number of scanned points is available, Random Forest would be the best suited algorithm for computing chemical element maps of paintings from XRF data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was partially funded by: Spanish Ministry of Economy and Competitiveness (grant number PID2020.118638RB.I00); ERDF funds; Regional Government of Andalusia/Ministry of Economic Transformation, Industry, Knowledge and Universities (grant number P18-RT-2248); the Health Institute Carlos III/Spanish Ministry of Science, Innovation and Universities (grant number PI20/00711); and Spanish Ministry of Science and Innovation (grant number PID2019-107793 GB-I00). Funding for publication charges: Universidad de Granada - Spain.

References

[1] C. Vanhoof, J.R. Bacon, U.E.A. Fittschen, L. Vincze, Atomic spectrometry update – a review of advances in X-ray fluorescence spectrometry and its special applications, J. Anal. At. Spectrom. 35 (2020) (2020) 1704–1719, https://doi.org/10.1039/D0JA90051F.

[2] A. Bezur, L.F. Lee, M. Loubser, K. Trentelman, Handheld XRF in cultural heritage: a Practical workbook for conservators, getty conservation Institute. https://books.google.it/books?id=m30uwwEACAAJ, 2019.

[3] J.D. Martin-Ramos, G. Chiari, SmART_scan: a method to produce composition maps using any elemental, molecular and image data, J. Cult. Herit. 39 (2019) 260–269, https://doi.org/10.1016/j.culher.2019.04.003.

[4] D. Miriello, R. De Luca, A. Bloise, G. Niceforo, J.D. Martin-Ramos, A. Martellone, B. De Nigris, M. Osanna, G. Chiari, Pigments mapping on two mural paintings of the "house of garden" in pompeii (campania, Italy), Mediterr. Archaeol. Archaeom. 21 (2021) 257–271, https://doi.org/10.5281/zenodo.4574643.

[5] E. Manzano, R. Blanc, J.D. Martin-Ramos, G. Chiari, P. Sarrazin, J.L. Vilchez, A combination of invasive and non-invasive techniques for the study of the palette and painting structure of a copy of Raphael's Transfiguration of Christ, Herit. Sci. 9 (2021) 150, https://doi.org/10.1186/s40494-021-00623-z.
[6] G. Chirco, M. de Cesare, G. Chiari, S. Maaß, M.L. Saladino, D.F. Chillura Martino, Archaeometric study of execution techniques of white Attic vases: the case of the Perseus crater in Agrigento, R. Soc. Chem. Adv 12 (2022) 4526–4535, https://doi.org/10.1039/D1RA06453C.
[7] G. Chirco, G. Chiari, D.F.C. Martino, Processing of XRF elementary data from the painted ceramic surface with innovative tools, J. Phys. Conf. Ser. 2204 (2022) 12083, https://doi.org/10.1088/1742-6596/2204/1/012083.
[8] T. Kleynhans, C.M. Schmidt Patterson, K.A. Dooley, D.W. Messinger, J.K. Delaney, An alternative approach to mapping pigments in paintings with hyperspectral reflectance image cubes using artificial intelligence, Herit. Sci. 8 (2020) 1–16, https://doi.org/10.1186/S40494-020-00427-7/FIGURES/6.
[9] A. Chen, R. Jesus, M. Vilarigues, Convolutional neural network-based pure paint pigment identification using hyperspectral images, ACM Int. Conf. Proceeding Ser. (2021), https://doi.org/10.1145/3469877.3495641.
[10] C. Jones, N.S. Daly, C. Higgitt, M.R.D. Rodrigues, Neural network-based classification of X-ray fluorescence spectra of artists' pigments: an approach leveraging a synthetic dataset created using the fundamental parameters method, Herit. Sci. 10 (2022) 1–14, https://doi.org/10.1186/S40494-022-00716-3/TABLES/5.
[11] B.J. Xu, Y. Wu, P. Hao, M. Vermeulen, A. McGeachy, K. Smith, K. Eremin, G. Rayner, G. Verri, F. Willomitzer, M. Alfeld, J. Tumblin, A. Katsaggelos, M. Walton, Can deep learning assist automatic identification of layered pigments from XRF data? J. Anal. At. Spectrom. 37 (2022) 2672–2682, https://doi.org/10.1039/D2JA00246A.
[12] J. Aldrich, Fisher and regression, stat, Sci. 20 (2005) 401–417, https://doi.org/10.1214/088342305000000331.
[13] J. Han, M. Kamber, J. Pei, Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, Waltham, Mass, 2012 https://doi.org/10.1016/C2009-0-61819-5.
[14] M.J. Gacto, J.M. Soto-Hidalgo, J. Alcalá-Fdez, R. Alcalá, Experimental study on 164 algorithms available in software tools for solving standard non-linear regression problems, IEEE Access 7 (2019) 108916–108939, https://doi.org/10.1109/ACCESS.2019.2933261.
[15] N. Matloff, Statistical Regression and Classification. From Linear Models to Machine Learning, Chapman and Hall/CRC, New York, 2017 https://doi.org/10.1201/9781315119588.
[16] Raphael, The transfiguration. https://www.museivaticani.va/content/museivaticani/en/collezioni/musei/la-pinacoteca/sala-viii—secolo-xvi/raffaello-sanzio–trasfigurazione.html, 2024.
[17] Niton, Niton XL2. https://www.thermofisher.com/order/catalog/product/XL2, 2022.
[18] Unknown, The miraculous interventions of Jizō Bosatsu. https://asia.si.edu/object/F1907.375a/, 2024.
[19] K.L. Rowberg, G. Hystad, M.L. Clarke, J. Gonzalez, J.M. Taylor, Mixing chemistry and pigments: X-ray fluorescence spectroscopy as a nondestructive technique for analysis of pigments in a painted Japanese handscroll, in: Context. Chem. Art Archaeol. Inspir. Instr., American Chemical Society, 2021, pp. 10–217, https://doi.org/10.1021/bk-2021-1386.ch010.
[20] M.L. Clarke, F. Gabrieli, K.L. Rowberg, A. Hare, J. Ueda, B. McCarthy, J.K. Delaney, Imaging spectroscopies to characterize a 13th century Japanese handscroll, the miraculous interventions of Jizō Bosatsu, Herit. Sci. 9 (2021) 20, https://doi.org/10.1186/s40494-021-00497-1.
[21] S. Weisberg, Applied Linear Regression, Wiley, 2005 https://doi.org/10.1002/0471704091.
[22] E. Ostertagová, Modelling using polynomial regression, Procedia Eng. 48 (2012) 500–506, https://doi.org/10.1016/j.proeng.2012.09.545.
[23] L. Breiman, Random Forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[24] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (2001) 1189–1232, https://doi.org/10.1214/aos/1013203451.
[25] J.R. Quinlan, Combining instance-based and model-based learning, in: Proc. Tenth Int. Conf. Int. Conf. Mach. Learn., Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993, pp. 236–243.
[26] M. Riedmiller, Advanced supervised learning in multi-layer perceptrons — from backpropagation to adaptive learning algorithms, Comput. Stand. Interfaces 16 (1994) 265–278, https://doi.org/10.1016/0920-5489(94)90017-5.
[27] C.-C. Chang, C.-J. Lin, LIBSVM: a library for Support vector machines, ACM Trans. Intell. Syst. Technol. 2 (2011), https://doi.org/10.1145/1961189.1961199.
[28] D.W. Aha, D. Kibler, M.K. Albert, Instance-based learning algorithms, Mach. Learn. 6 (1991) 37–66, https://doi.org/10.1007/BF00153759.
[29] M.R. Berthold, N. Cebron, F. Dill, T.R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel, KNIME: the konstanz information miner, stud. Classif. Data anal. Knowl, Organ (2008) 319–326, https://doi.org/10.1007/978-3-540-78246-9_38.
[30] S. Arlot, A. Celisse, A survey of cross-validation procedures for model selection, Stat. Surv. 4 (2010) 40–79, https://doi.org/10.1214/09-SS054.
[31] D.W. Gareth James Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning : with Applications in R, Springer, New York, 2021. https://search.library.wisc.edu/catalog/9910207152902121.
[32] G.C. Cawley, N.L.C. Talbot, On over-fitting in model selection and subsequent selection bias in performance evaluation, J. Mach. Learn. Res. 11 (2010) 2079–2107. http://jmlr.org/papers/v11/cawley10a.html.
[33] J. Demšar, Statistical comparisons of classifiers over multiple data sets, J. Mach. Learn. Res. 7 (2006) 1–30. http://jmlr.org/papers/v7/demsar06a.html.
[34] S. García, A. Fernández, J. Luengo, F. Herrera, Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: experimental analysis of power, Inf. Sci. 180 (2010) 2044–2064, https://doi.org/10.1016/j.ins.2009.12.010.
[35] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, J. Am. Stat. Assoc. 32 (1937) 675–701, https://doi.org/10.1080/01621459.1937.10503522.
[36] S. Holm, A simple sequentially rejective multiple test procedure, Scand. J. Stat. 6 (1979) 65–70. http://www.jstor.org/stable/4615733.
[37] S. Raschka, J. Patterson, C. Nolet, Machine learning in Python: main developments and technology trends in data science, Machine Learning, and Artificial Intelligence, Information 11 (2020), https://doi.org/10.3390/info11040193.