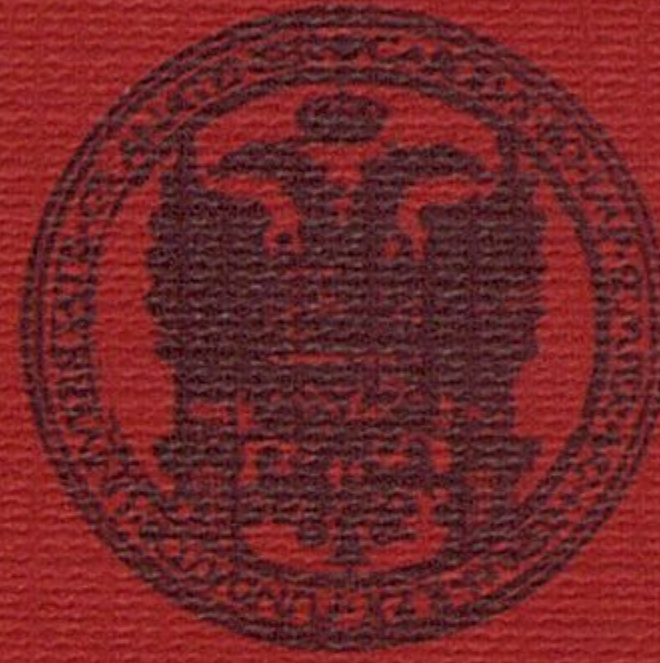


UNIVERSIDAD DE GRANADA  
FACULTAD DE CIENCIAS



CLASIFICACIÓN DE LESIONES MAMARIAS  
MALIGNAS VERSUS BENIGNAS MEDIANTE  
ANÁLISIS DE IMAGEN APLICADO A  
ESPECÍMENES CITOLÓGICOS OBTENIDOS  
POR PUNCIÓN ASPIRACIÓN

TESIS DOCTORAL

CARMEN HERRERA ESPÍNEIRA



Biblioteca Universitaria de Granada



01533854

GRANADA 1995



PROV. T-15/96

UNIVERSIDAD DE GRANADA  
Facultad de Ciencias  
Fecha .....5-6-95.....  
ENTRADA NUM. ....1267.....



T  
15  
70

---

UNIVERSIDAD DE GRANADA


FACULTAD DE CIENCIAS

TITULO:

CLASIFICACIÓN DE LESIONES MAMARIAS MALIGNAS VERSUS BENIGNAS  
MEDIANTE ANÁLISIS DE IMAGEN APLICADO A ESPECÍMENES CITOLÓGICOS  
OBTENIDOS POR PUNCIÓN ASPIRACIÓN.

BIBLIOTECA UNIVERSITARIA	
GRANADA	
Nº Documento	519676724
Nº Copia	21223403

Tesis presentada por Carmen Herrera Espiñeira,  
licenciada en Ciencias (sección Biológicas) por la  
Universidad de Granada, para optar al grado de Doctora.



Granada 1995



---

D. Jose Javier Esquivias López-Cuervo, Doctor por la Facultad de Medicina de Granada, D. Joaquín Fernández Valdivia, Doctor por la Escuela Técnica Superior de Ingeniería Informática y D. Jose Luis Caballero López-Lendinez, Doctor por la Facultad de Ciencias de Granada (sección Biológicas)

CERTIFICAN: Que la tesis doctoral que presenta, al superior juicio del tribunal, Dña. CARMEN HERRERA ESPIÑEIRA con el título:

**CLASIFICACIÓN DE LESIONES MAMARIAS MALIGNAS VERSUS BENIGNAS MEDIANTE ANÁLISIS DE IMAGEN APLICADO A ESPECÍMENES CITOLÓGICOS OBTENIDOS POR PUNCIÓN ASPIRACIÓN.**

ha sido realizada bajo nuestra dirección y por cuyo resultado la estimamos acreedora al título de Doctora, siempre que así lo estime el citado tribunal.

FIRMADO:

J. Javier Esquivias López-Cuervo

Joaquín Fernández Valdivia

J. Luis Caballero López-Lendinez



Granada a 30 de Abril de 1995



---

Este estudio se enmarca en el proyecto de investigación "CLASIFICACIÓN AUTOMÁTICA A PARTIR DE ANÁLISIS DE IMAGEN USANDO MAMOGRAFÍAS, CITOLOGÍAS Y CORTES HISTOLÓGICOS" del Programa Sectorial de Promoción General del Conocimiento de la Dirección General de Investigación Científica y Técnica (DGICYT) con el número de referencia: PM90-0093-C02-01. Servicio Andaluz de Salud. Organismo Ejecutor: Hospital General de Especialidades "Virgen de las Nieves" de Granada, Departamento de Anatomía Patológica.



---

## AGRADECIMIENTO

A J. Javier Esquivias López-Cuervo, mi mentor en análisis de imagen, impulsor y director principal de la memoria a las que tantas horas dedicó. Sin su ayuda éste trabajo no existiría.

A Joaquín Fernández Valdivia, director de ésta memoria, por la gran ayuda, siempre desinteresada, que sus conocimientos en Inteligencia Artificial supuso en la difícil comprensión de éste área, desde mucho antes de ser director de ella, y por su enorme paciencia para trasmitirmela. Sin él el resultado no habría sido el mismo.

A Jose Luis Caballero López-Lendinez, director de ésta memoria, por sus imprescindibles directrices en la forma y el contenido y su afable trato.

A D. Enrique Montero García, jefe del servicio de Anatomía Patológica del Hospital del SAS " Virgen de las Nieves de Granada". Promotor de ésta tesis a causa de su innato interés en fomentar la investigación, y por poner a mi disposición los recursos y materiales del servicio necesarios para el desarrollo del trabajo.

A Dña. Cristina Marcos Muñoz, jefe de sección del mismo servicio, notable citóloga, en cuyos diagnósticos y criterios citológicos se ha basado el proyecto de investigación en el que está enmarcada esta tesis, y por animar a su realización.

A D. Luis Rodríguez Rico, jefe de sección del mismo servicio, por su inagotable capacidad para enseñar, sólo debido a su enorme generosidad y su innata disposición para transmitir conocimientos.

A D. Alfonso Herruzo Nalda, jefe del departamento de Obstetricia y Ginecología del mismo hospital, por la información inicial que facilitó la localización de los casos objetos de estudio para el proyecto de investigación en el que está enmarcada ésta tesis.

A D. Antonio Rios Guadix, catedrático de Biología Celular y director del departamento de Biología Celular. Su gran disponibilidad y generosidad ha facilitado el que ésta tesis pueda presentarse.

A Laureano Garrido Pérez y Jesús Calle Romero por sus ayudas en diversas fases del trabajo y su imprescindible apoyo moral.

A mis compañeros del Servicio de Anatomía Patológica, porque a lo largo de tantos años en que se ha desarrollado el estudio han tenido que sobrellevarme en los momentos de inquietud.



---

A mi madre Dolores, por el pasado.

A mis hijos Sara y David, por el futuro.



---

## INDICE

JUSTIFICACION DEL TRABAJO.....pag.1

### I - INTRODUCCION

I-1 ANATOMIA MICROSCOPICA.....4

I-2 ANATOMIA PATOLOGICA .....5

I-2-1 Lesiones benignas de la mama.....8

I-2-2 Lesiones malignas de la mama.....9

I-2-3 Clasificación de las lesiones mamarias.....10

I-3 EL CARCINOMA DE MAMA.....13

I-3-1 Incidencia.....13

I-3-2 Factores Etiológicos.....14

I-3-3 Factores Pronósticos.....15

I-3-3-A Extensión del tumor: Clasificación TNM.....16

I-3-3-B Factores cito-histológicos.....17

a-Índice de Scarff, Bloom, y Richardson (SBR)

b-Actividad proliferativa

I-3-3-C Factores Endocrinos: Receptores hormonales.....20

I-3-3-D Factores Genéticos: Expresión de Oncogenes.....21

I-4 METODOS DE ESTUDIO CONVENCIONALES DE LA PATOLOGIA MAMARIA

I-4-1 Exploración Clínica.....23

I-4-2 Mamografía.....23

I-4-3 Ecografía.....24

I-4-4 Citología por Punción Aspiración.....25

I-4-4-1 Criterios de malignidad citológicos.....26

I-5 METODOS DE ESTUDIO NO CONVENCIONALES:

Métodos Cuantimétricos.....31

I-5-1 Métodos de estudio Fotométricos o Densitométricos.31

I-5-1-A Microespectrometría o Citofotometría.....31

I-5-1-B Citometría de Flujo.....33

I-5-2 Métodos de estudio Morfométricos.....33

I-5-3 Análisis Textural.....34

I-5-4 Análisis Computerizado de la Imagen.....37

I-5-4-1 Metodología.....38

a-Captación de la imagen

b-Procesamiento de la imagen

c-Análisis de la imagen



---

I-6 METODOS DE CLASIFICACION DE DATOS.....	44
I-6-1 Taxonomía de los métodos de clasificación.....	44
I-6-1-A Aproximaciones Estadísticas.....	44
I-6-2-B Aprendizaje.....	46
I-6-2-C Redes Neuronales.....	47
I-7 OBJETIVOS DEL ESTUDIO.....	49
II MATERIAL Y METODO.....	50
II-1 PROCESO DE CALIBRADO DE LA CAMARA.....	57
II-2 PROCESO DE CAPTACION DE LAS IMAGENES.....	58
II-2-1 Cámara de video.....	58
II-2-2 Depuración del proceso de captación.....	59
II-3 PROCESO DE SELECCION DE LAS IMAGENES.....	63
II-4 SEGMENTACION DE LA ZONA DE INTERES.....	64
II-4-1 Macro para la relación AN/AC en los grupos.....	64
II-4-2 Descripción de las funciones del software específico de análisis automático de los grupos basado en distancias nucleares: Creación de grafos planares.....	69
II-4-2 Descripción de las funciones del software específico de análisis automático para la segmentación automática de los núcleos y grupos grandes.....	72
a-Segmentación para casos de no solapamientos nucleares	
b-Segmentación para casos de solapamientos nucleares	
II-5 EL ANALISIS DE LOS DATOS.....	82
II-5-1 La Clasificación.....	83
II-5-1-A En base a la planimetría nuclear.....	84
II-5-1-B En base a texturas nucleares.....	84
a-Medidas basadas en el histograma de niveles de gris	
b-Medidas basadas en diferencias de niveles de gris	
c-Medidas basadas en coocurrencias de niveles de gris	
II-5-1-C En base a distancias nucleares en los grupos: Creación de grafos planares.....	88
II-5-1-D En base a la relación AN/AC en los grupos.....	89

---



---

<b>III RESULTADOS</b> .....	91
<b>III-1 RESULTADOS DEL ANALISIS DE LOS GRUPOS</b> .....	92
III-1-1 Experimento 1: Variables Planimétricas.....	92
III-1-1-A Resultados por casos con CART.....	96
III-1-1-B Resultados por imágenes con CART.....	96
III-1-1-C Resultados por casos con Análisis Discriminante.....	98
III-1-2 Experimento 2: Variables Texturales.....	100
III-1-2-A Resultados por casos con CART.....	100
III-1-2-B Resultados por imágenes con CART.....	103
III-1-2-C Resultados por casos con Análisis Discriminante.....	106
III-1-3 Experimento 3: Todas la Variables.....	107
<b>III-2 RESULTADOS DEL ANALISIS DE LOS GRUPOS</b> .....	108
III-2 Análisis de las distancias nucleares con grafos planares (medidas de desorden).....	108
III-2-A Resultados por casos con CART.....	108
III-2-B Resultados por grupos individuales con CART....	111
III-2-C Resultados por casos con Análisis Discriminante	113
<b>IV DISCUSION</b> .....	116
IV-1 CONSTRUCCION DEL CLASIFICADOR.....	117
IV-1-1 Selección de los rasgos citológicos.....	117
IV-1-2 Tratamiento del material par el análisis de imagen.....	121
IV-1-3 Metodología para el análisis de imágenes.....	122
IV-2 EL CLASIFICADOR.....	123
IV-2-1 La base de datos.....	124
IV-2-2 La segmentación automática y semiautomática....	125
IV-2-3 Las variables tratadas.....	127
IV-2-3 La metodología para el análisis de los datos....	136
<b>V LINEAS FUTURAS</b> .....	138
<b>VI CONCLUSIONES</b> .....	139
<b>VII BIBLIOGRAFIA</b> .....	141

---



---

## JUSTIFICACION DEL TRABAJO

El diagnóstico citológico de las lesiones mamarias lo realiza el experto analizando en la extensión citológica un conjunto de rasgos celulares mediante un mecanismo de patrones de reconocimiento adquiridos a través de la experiencia propia y ajena. Este análisis, que es fundamentalmente cualitativo y laborioso, está jerarquizado en una serie de variables biológicas que tienen diferente importancia en la clasificación de las lesiones. Encontrar variables discriminantes fácilmente cuantificables podría ser de valiosa ayuda diagnóstica.

Se pretende efectuar una cuantificación, con técnicas de análisis, de imagen del mayor número de variables posibles que el citólogo utiliza para llegar al diagnóstico, con objeto de conseguir un clasificador que divida las extensiones citológicas en benignas y malignas de forma rápida, eficaz y reproducible. Estas técnicas, que hasta no hace mucho eran difíciles y laboriosas, se han agilizado en extremo con las técnicas de Análisis de Imagen que aquí utilizaremos, con objeto de automatizar al máximo el proceso. Para ello se utilizarán extensiones citológicas de mama teñidas con la técnica habitual de forma que, realizar la cuantificación de las variables o aplicar el clasificador, sea un procedimiento rápido y eficaz de ayuda en la rutina diagnóstica.



---

En la exposición de ésta tesis doctoral se desarrollará en la Introducción, en primer lugar y de forma sucinta, la anatomía microscópica de la mama, su patología, con especial referencia al carcinoma, y tras enumerar los métodos de estudio de la patología mamaria, haremos descripción más detallada de los rasgos citológicos que permiten realizar el diagnóstico. A continuación se comentaran los métodos de análisis de imagen y su aplicación a la citología mamaria, así como los métodos de clasificación.

En el apartado de Material y Métodos se hará referencia a los casos estudiados y se explicará la metodología seguida que, en esencia, consiste en valorar los criterios citológicos que llevan a la catalogación de las lesiones como benignas y malignas en base a LOS RASGOS GENERALES DEL EXTENDIDO, RASGOS GENERALES DE LOS GRUPOS DE CÉLULAS Y RASGOS ESPECÍFICOS DE LOS NÚCLEOS CELULARES (Morfométricos y Texturales). Tras ensayar todas estas variables se seleccionarán las que sean susceptibles de ser sometidas a análisis de imagen, intentando siempre reproducir la sistemática que se emplea en la práctica habitual y procurando que la segmentación de la imagen pueda llevarse a cabo de la manera más automatizada posible.

La mayoría de los clasificadores analizan y seleccionan los rasgos mediante clasificadores estadísticos, generalmente a través de un análisis multivariante. Existe,



---

sin embargo, un procedimiento que se adapta muy bien a la forma en que el experto realiza el diagnóstico: son los árboles de clasificación, mecanismo que se ha mostrado como uno de los fundamentales para la selección y el análisis de variables estimadas, vía técnicas de análisis de imagen. En patología mamaria se ha aplicado este método de análisis de datos con variables cualitativas.

En el apartado de **Resultados** se dará cuenta de lo obtenido con la aplicación de los algoritmos concebidos expresamente para realizar la segmentación de las imágenes captadas a partir de las extensiones citológicas previamente seleccionadas, y también de la Sensibilidad y Especificidad del método de clasificación, tras realizar el análisis de imagen.

En la **Discusión** se practicará una exposición crítica de los resultados obtenidos y se compararán éstos con los obtenidos por otros autores con distintos o semejantes procedimientos.



---

## INTRODUCCION



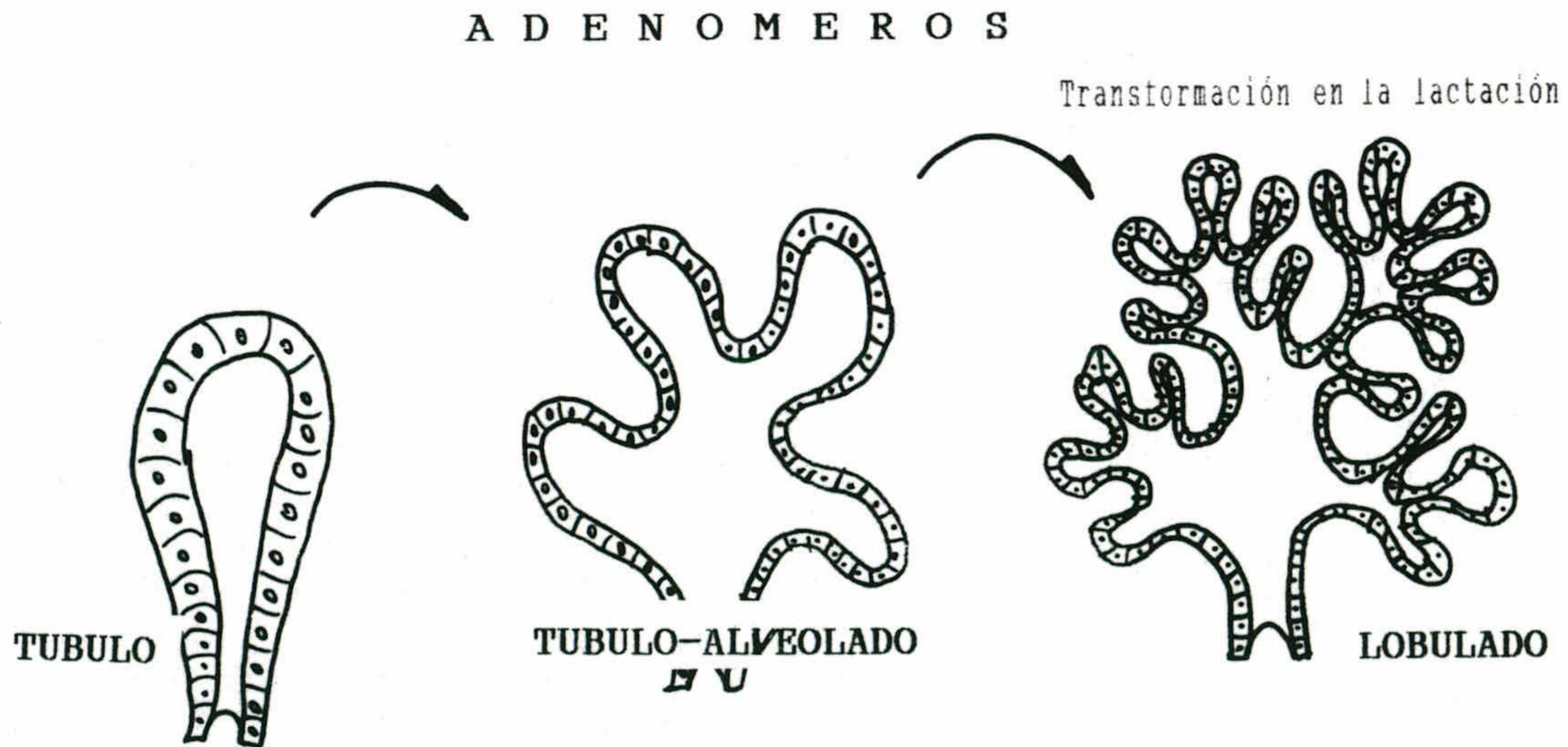
---

# I-INTRODUCCION

## I-1 ANATOMIA MICROSCOPICA

La glándula mamaria es una glándula especializada de la piel. Está formada por 15 o 20 lóbulos subdividido en lobulillos, cada uno de éstos es una glándula compuesta por una masa de adenómeros (porción secretora de la glándula), que pueden ser alveolos o túbulos. (Figura 1).

Figura 1: Dibujo modificado de Russo and Russo 1987





---

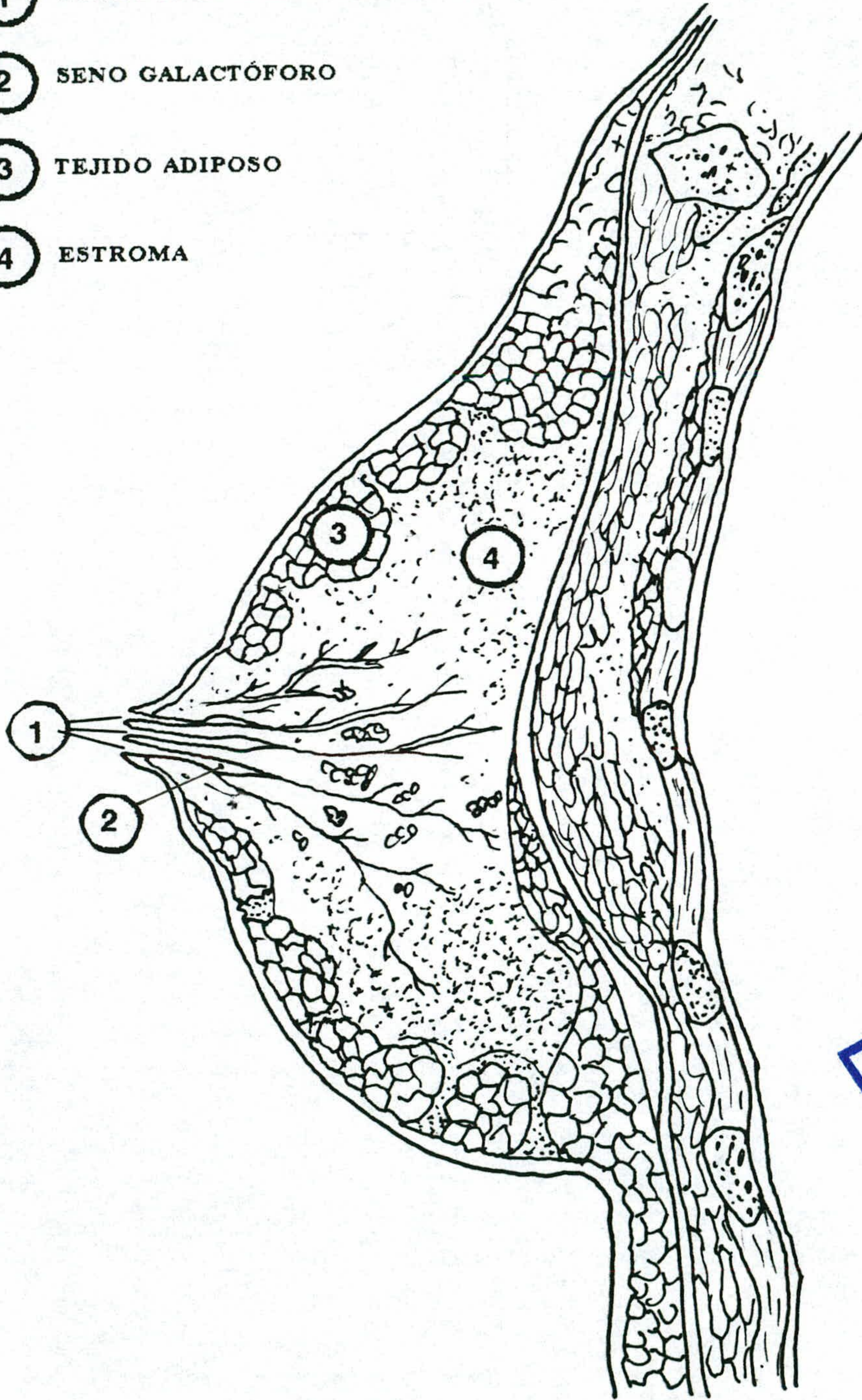
El estroma es de tipo tanto fibroso como adiposo. Cada lóbulo es una glándula compuesta independiente cuyos conductos primarios se reúnen una y otra vez para formar conductos mayores que desembocan en un conducto galactóforo único de 2 a 4,5 mm de diámetro que, antes de desembocar en el pezón, presenta una dilatación llamada seno lactífero. El epitelio de revestimiento del conducto galactóforo a nivel del pezón empieza siendo plano estratificado y presenta una transición gradual hasta pasar a una única capa de células cilíndricas en el resto del sistema ductal. Las porciones secretoras de la glándula (conductos alveolares y alveolos), están formados por células secretoras cuboideas que reposan sobre una lámina basal y sobre las prolongaciones subyacentes de las células mioepiteliales, éstas envuelven a los alveolos en una red a modo de malla (Azzopardi 1979, Fawcett 1989). (Figura 2).



Figura 2

**LEYENDA**

- ① CONDUCTOS GALACTOFÓRICOS
- ② SENO GALACTÓFORO
- ③ TEJIDO ADIPOSO
- ④ ESTROMA



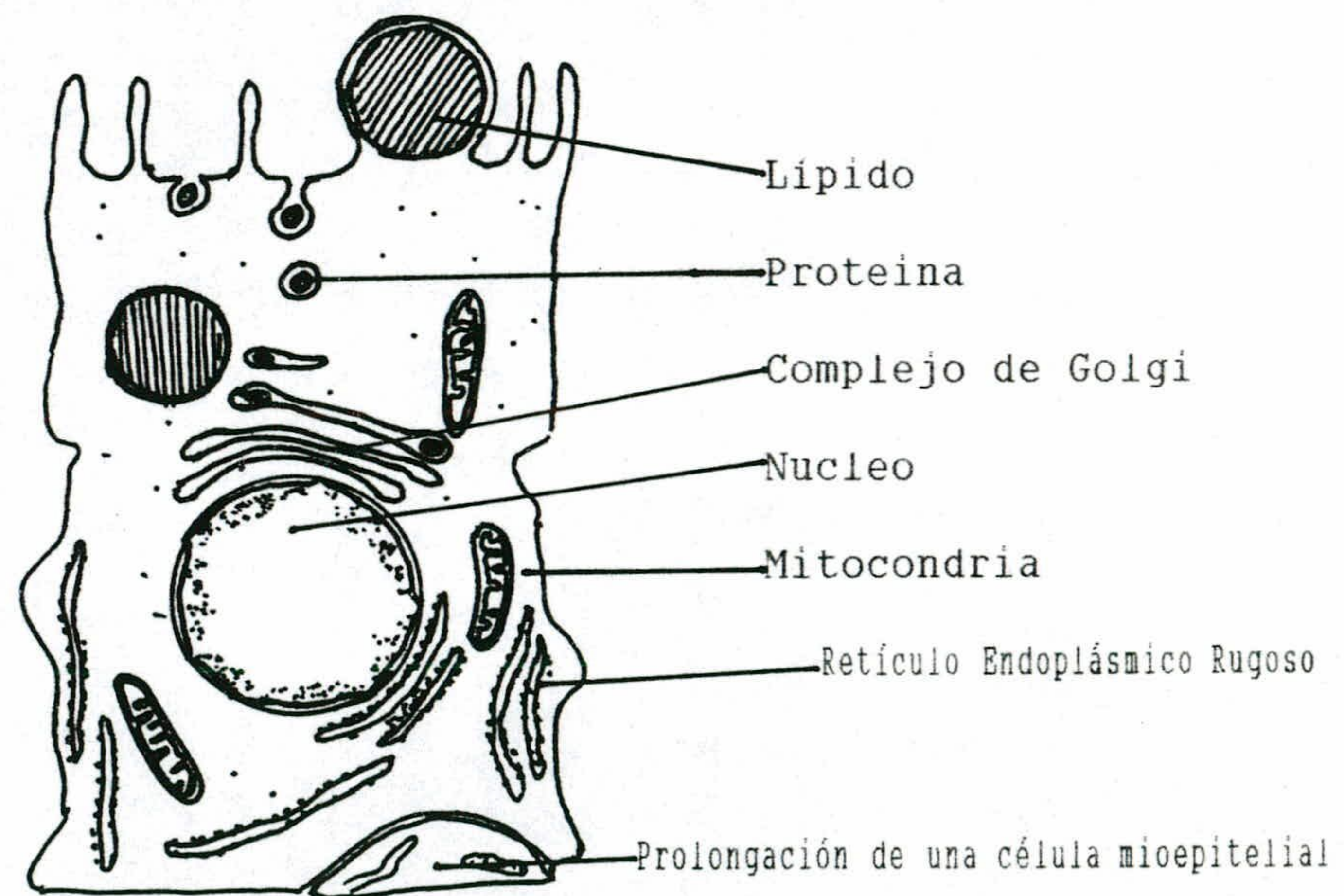
UNIVERSIDAD DE GRANADA  
24 MAYO 1995  
COMISION DE DOCTORAL



---

La célula epitelial secretora es fundamentalmente acidófila, varía de plana a cilíndrica, con núcleo redondo u oval, situado en la zona media de la célula. Segrega dos productos secretorios distintos: protéico y lipídico. Los componentes proteicos se vierten tras fusionarse con la membrana citoplasmática (liberación merocrina), sin embargo el otro producto secretorio: los lípidos son liberados junto con una porción de la membrana celular (liberación apocrina) (Fawcett 1989, Maillet 1990). (Figura 3).

Figura 3: CÉLULA EPITELIAL SECRETORA MAMARIA



Modificado de Fawcett 1989. Dibujo basado en las observaciones de Bargmann W, y A Koop 1959



---

Existen otros tipos de células en la glándula mamaria, en el tejido conjuntivo que la rodea están presentes dos tipos de poblaciones celulares: una población relativamente estable de células fijas como los fibroblastos (responsable de la producción de la matriz extracelular) y células adiposas (que almacenan material energético de reserva), y una población de células emigrantes: macrófagos, eosinófilos, linfocitos y células cebadas.

## **I-2 ANATOMIA PATOLOGICA**

Las lesiones proliferativas mamarias se presentan con una extraordinaria frecuencia entre la población femenina, yendo progresivamente en aumento con la edad. Estas lesiones pueden ser malignas o benignas.

### **1-2-1 LESIONES BENIGNAS DE LA MAMA**

Entre las lesiones benignas mamarias son más frecuentes el Fibroadenoma y la Enfermedad fibroquística.

El FIBROADENOMA se presenta como una tumoración variable (entre 1 y 5 cm), en general única y a veces múltiple, compuesta de epitelio derivado de los pequeños conductos y tejido conectivo del estroma intralobulillar (Koss 1988).



---

LA ENFERMEDAD FIBROQUÍSTICA es una combinación de alteraciones proliferativas y regresivas generalizada de los tejidos mamarios que se manifiesta por hiperplasias epiteliales, quistes y fibrosis alrededor de los conductos. La hiperplasia, dependiendo de la zona de la glándula en que se localice, produce un tipo de lesión histológica diferente. Así, si las proliferaciones son extralobulillares con aumento intracanalicular del epitelio, se denomina Hiperplasia Canalicular o Epiteliosis. Si las proliferaciones son intralobulillares con aumento intracanalicular del epitelio y proliferaciones de los canaliculos intralobulillares se denomina Hiperplasia Lobulillar o Adenosis; los Quistes suelen ser múltiples y frecuentemente revestidos de células apocrinas, epitelio atrófico o hiperplasias papilares (OMS 1982).

### 1-2-2 LESIONES MALIGNAS DE LA MAMA

El cáncer es una tumoración que se caracteriza por el crecimiento anárquico de sus células, la capacidad de emigración de éstas para producir metástasis y capacidad para eludir la vigilancia inmunológica (Santos y Villanueva 1985). Esta capacidad migratoria los distingue de los tumores benignos (Cairns 1985).

La degeneración neoplásica maligna del componente epitelial de la glándula mamaria origina el carcinoma de



---

mama, mientras que las neoformaciones que surgen del estroma constituyen los sarcomas. Entre las lesiones malignas mamarias es el CARCINOMA CANALICULAR O DUCTAL el más frecuente constituyendo del 75 al 85 % de todos los cánceres mamarios en la mujer (Koss 1988). De éste porcentaje sólo del 3 al 5 % no son invasores (Ayala 1988), siendo su localización más frecuente los cuadrantes superiores por ser los que tienen mayor volumen (Uriburu 1977 y Fernandez-Cid 1982); éste tipo de tumor puede presentar diversos patrones histológicos según la cantidad de estroma fibroso o según el tipo estructural (OMS 1982). (Figura 4a y 4b).

### **1-2-3 CLASIFICACION DE LAS LESIONES MAMARIAS**

Atendiendo a la clasificación histológica de la Organización Mundial de la Salud (OMS 1982), las lesiones de la mama se dividen en las distintas clases que figuran en la tabla 1.



---

Figura 4a: CORTE HISTOLÓGICO DE LESIÓN MAMARIA BENIGNA  
(imagen tomada a 100 aumentos)

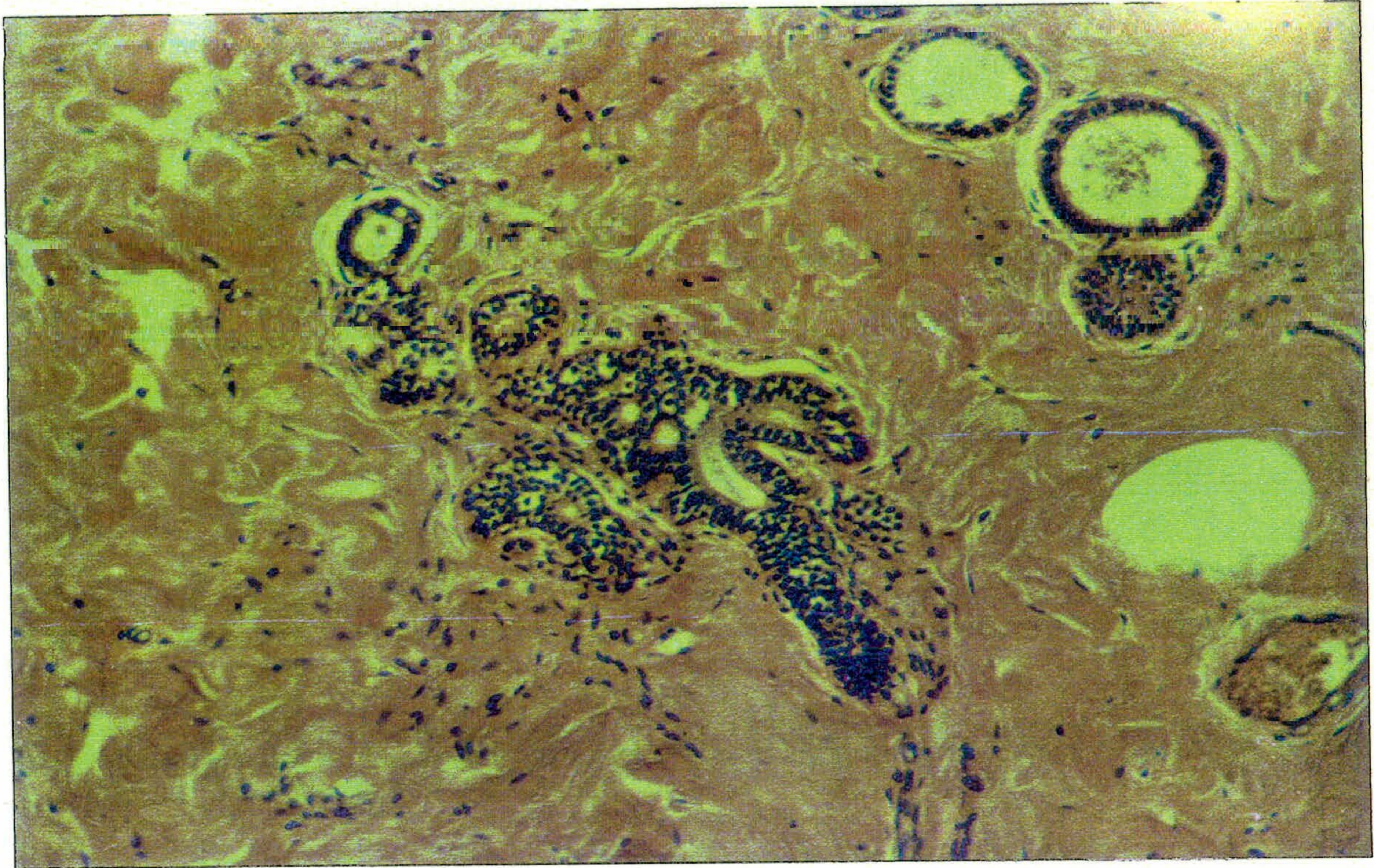


Figura 4b: CORTE HISTOLÓGICO DE LESIÓN MAMARIA MALIGNA  
(imagen tomada a 100 aumentos)

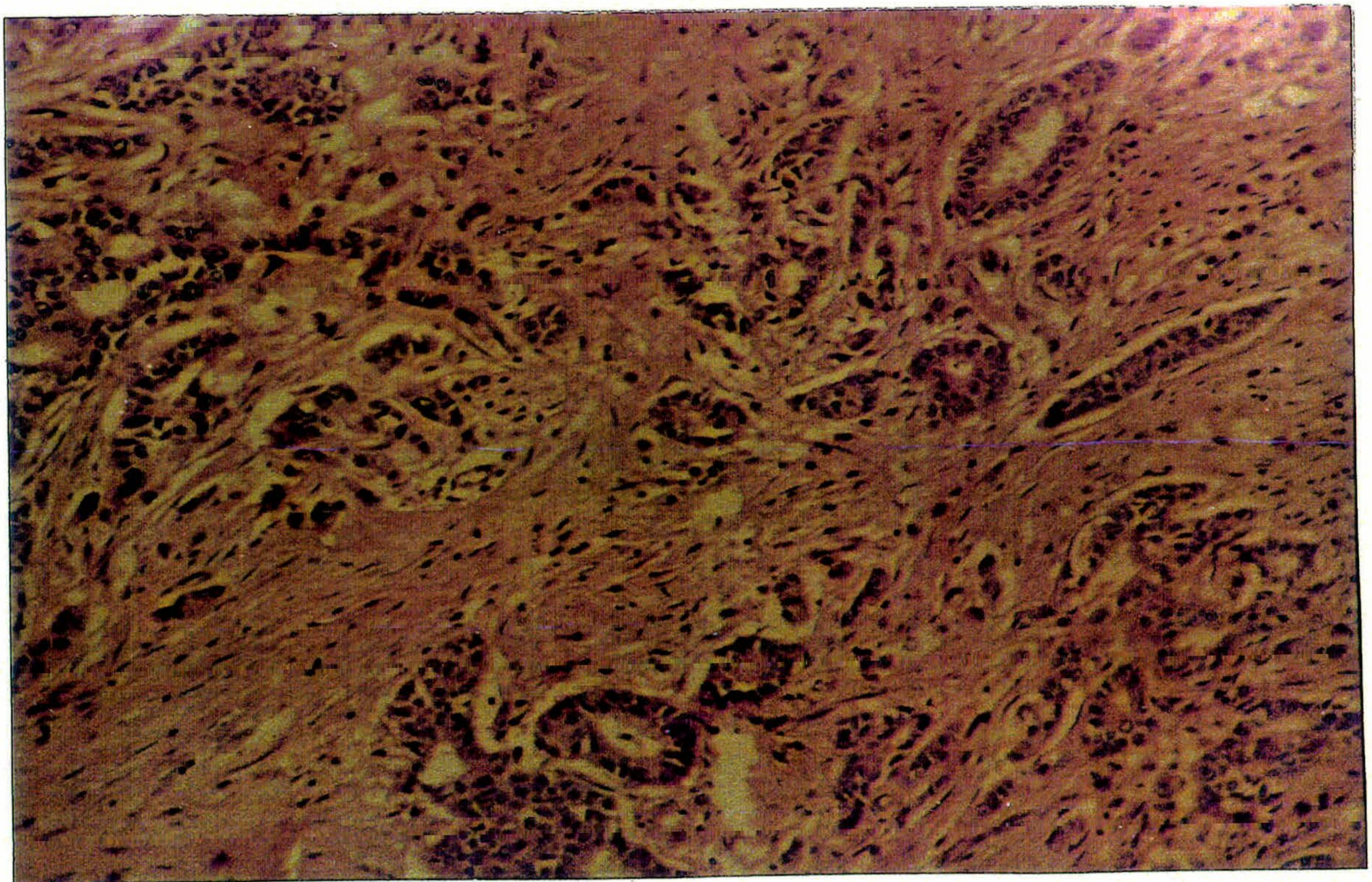




Tabla 1

<p><u>a- Tumores Epiteliales Benignos:</u> Papiloma intracanalicular Adenoma del pezón Adenoma tubular y lactante</p>
<p><u>b- Tumores Epiteliales Malignos:</u></p> <ul style="list-style-type: none"><li>- No invasores: Carcinoma intracanalicular Carcinoma lobulillar</li><li>- Invasores: Carcinoma lobulillar Carcinoma canalicular Canalicular con predominio intracanalicular Carcinoma mucinoso Carcinoma medular Carcinoma papilar Carcinoma tubular Carcinoma Adenoide-quístico Carcinoma secretor juvenil Carcinoma apocrino Carcinoma con metaplasia</li><li>- Otros: Liposecretor Inflamatorio Enfermedad de Paget del pezón</li></ul>
<p><u>c- Tumores Mixtos de tejido conjuntivo y epitelio</u> Fibroadenoma (benigno) Tumor Filodes (benigno y maligno) Carcinosarcoma</p>
<p><u>d- Tumores varios</u></p> <ul style="list-style-type: none"><li>- de los tejidos blandos: Angiosarcoma Hemangioma Sarcoma estromático Mioblastoma</li><li>- de la piel</li><li>- de los tejidos hematopoyéticos y linfoides</li></ul>
<p><u>e- Displasia Mamaria o Enfermedad Fibroquística</u></p>
<p><u>f- Otros no clasificables</u></p>
<p><u>g- Lesiones seudotumorales</u> Ectasia canalicular Seudotumores inflamatorios: - Hamartoma - Ginecomastia</p>



---

## I-3 EL CARCINOMA DE MAMA

### I-3-1 INCIDENCIA

La mama es una de los primeros lugares de localización de cáncer en la mujer en los países desarrollados (menos del 1% de los cánceres mamarios afectan al hombre). Su mortalidad incluso ha aumentado en los últimos 30 años, siendo rara su aparición por debajo de los 25 años y aumentando progresivamente a partir de esta edad hasta alcanzar su máximo entre los 45 y 60 años (Uriburu 1977). A lo largo de éste periodo se detectan dos picos máximos de incidencia alrededor de los 45 y 65 años (De Waard 1964, Herrera 1993). Todo parece indicar, como se ilustra posteriormente, que el primero de ellos está más inluido por factores genéticos y el segundo por factores ambientales.

Aproximadamente el 10 % de los carcinomas mamarios ya tienen metástasis en el momento de su diagnóstico, siendo la supervivencia sin evidencia de metástasis a los 6,5 años de aproximadamente el 46%, de los cuales el 34% presentaban ya metástasis ganglionares en el momento de la extirpación quirúrgica (Baak 1982).

En España las provincias de alto riesgo son: Gerona, Baleares, Tarragona, Zaragoza y Barcelona, las de bajo riesgo: Guadalajara, Lugo, Jaén, Cuenca, Almería,



---

Granada (Atlas del Cáncer en España 1984). En Andalucía en general y Granada en particular, ocupa el segundo lugar de muerte por cáncer en la mujer, representando en Andalucía el 13 % del total de muertes (Mortalidad en Andalucía 1987).

### **1-3-2 FACTORES ETIOLOGICOS**

El cáncer de mama parece una enfermedad multifactorial. En su desarrollo se barajan factores genéticos, socio-culturales, hormonales, víricos...etc. (Gallagher 1969, Charney 1971, Lynch 1989).

Entre los factores de riesgo que más frecuentemente se barajan están, con riesgo moderado, el padecer alguna hiperplasia mamaria (Uriburu 1977), aumentando considerablemente cuando ésta hiperplasia es atípica (Consensus Meeting 1986, Dupont 1985), la raza blanca, la posición económica alta y el vivir en el hemisferio occidental (Uriburu 1977). Sin embargo la tasa de incidencia de cáncer de mama en los emigrantes se va haciendo progresivamente equiparable al del país de adopción lo que sugiere que son fundamentalmente las causas ambientales las que establecen las diferencias (Thomas 1978, Henderson 1984). Entre éstas causas ambientales se destaca la dieta rica en grasas (Thomas 1978, Hirohata 1987), fundamentalmente en lípidos poliinsaturados (Gammal 1967, Carroll 1971 y 1979), el sobrepeso y alimentos ricos en colesterol (Wynder 1980 y 1984, Howson 1986). Otros factores como



---

la exposición a radiaciones, el consumo de alcohol, la paridad, el uso de anticonceptivos orales o la lactancia, son más controvertidos (Domochowski 1969, Ing 1977, Haagensen 1981, Levit 1983, Pike 1983, Rosner 1985, Kvale 1987, Schatzkin 1987, Willet 1987, Hrubec 1989, Anderson 1989).

Se sospechaba que existían factores genéticos que podían influir en la aparición del cáncer de mama (Thor 1986, Theillet 1986, Ferti-Passantonopoulou 1987). Los estudios con árboles genealógicos apuntan a una transmisión genética de tipo autosómico dominante que sería la causa de las formas heredables de la enfermedad (Lynch 1989 y 1994) que no representan más del 8 % de todos los carcinomas mamarios y que se caracteriza por: aparición temprana, alto riesgo de afectación bilateral y familiares femeninos con más riesgo de padecerla que el resto de la población (Cady 1970, Lynch 1989, Anderson 1989).

En general se acepta que el medio ambiente influye en la aparición del cáncer postmenopáusico y los factores genéticos en el premenopáusico (Atlas del Cáncer en España 1984).

### **I-3-3 FACTORES PRONOSTICOS**

Como factores que orientan sobre la evolución que va a tomar la enfermedad, la bibliografía es muy amplia al respecto. Entre los más importantes tenemos:



---

A- EXTENSIÓN DEL TUMOR.

B- FACTORES CITO-HISTOLÓGICOS (Índice de Scarff, Bloom, Richardson en histología y citología, actividad proliferativa celular, corpúsculo sexual, morfometría núcleo/citoplasmática).

C- FACTORES ENDOCRINOS (receptores hormonales).

D- FACTORES GENÉTICOS (activación de Oncogenes).

### **I-3-3-A Extensión del tumor: Clasificación**

#### **TNM**

Este sistema se basa en la predicción de la supervivencia por (UICC 1979):

- El tamaño del tumor primario (T)
- La afectación ganglionar regional (N)
- La ausencia o presencia de metástasis (M)

Russo y col. en 1987 concluyen de sus estudios que el estado de los ganglios linfáticos, el tamaño del tumor y el grado del tumor tenían un valor pronóstico significativo y, de ellos, la presencia o ausencia de ganglios axilares sería el parámetro predictivo más simple del desenlace de la enfermedad. El porcentaje libre de recurrencia a los cuatro o cinco años para las pacientes sin afectación ganglionar se situaría entre el 80 al 85% y del 55 al 62 % para aquellas con afectación ganglionar (Baak 1985, Linden 1987).



---

### **I-3-3-B Factores cito-histológicos:**

#### a- Índice de Scarff, Bloom y Richardson (SBR)

Se utiliza para clasificar los cánceres de mama en grados en base a:

- La tendencia a formar estructuras glandulares semejantes al tejido sano.
- Variación en el tamaño (pleomorfismo celular)
- Frecuencia de mitosis y picnosis celular

El valor pronóstico del método ha sido demostrado en los trabajos de Bloom 1957, Davis 1986 y Doussal 1989, entre otros.

En citología el Índice de Scarff, Bloom y Richardson se basa en la adaptación de los tres factores valorados en la histología a los extendidos citológicos (Zajdela 1981, Marcos 1986).

#### b- Actividad proliferativa

Se puede estimar de varias formas: mediante el ÍNDICE MITÓTICO, CUANTIFICACIÓN DEL DNA, VALORACIÓN DE LOS CORPÚSCULOS SEXUALES, VARIABLES MORFOMÉTRICAS NUCLEO/CITOPLASMÁTICAS Y OTROS MÉTODOS DE CUANTIFICACIÓN.



---

Baak en 1982 y 1985 encuentra que el **ÍNDICE MITÓTICO** es el más importante factor pronóstico en el cáncer de mama seguido por la cuantificación morfométrica de los parámetros nucleares y el grado histológico y nuclear. Otros trabajos también ratifican su valor pronóstico (Walgren 1976, Linden 1986).

LA CUANTIFICACIÓN DEL DNA se ha realizado en diversos tipos de tumores ya sea por técnicas de microespectrofotometría (citofotometría) (Adams 1968, Zetterberg 1976, Balaban-Malembaum 1979) o técnicas de citometría de Flujo (Kajikawa 1978, Ronstrom 1981, Diamomond 1982). En las neoplasias humanas es frecuente las anomalías en el contenido de DNA: del 70 % al 80 % de los tumores sólidos lo poseen y está en relación con la mayor indiferenciación de las células, la evolución y la actividad proliferativa del tumor (fracción de fase S, índice mitótico). Por ello es de gran utilidad la determinación del Índice de DNA (cociente entre el contenido de DNA del tumor y el contenido normal) (Brugarolas 1989). Incluso un mismo tumor puede tener gran heterogeneidad de poblaciones celulares que condicionan diferente comportamiento de sus células en el crecimiento, capacidad metastatizante o respuesta al tratamiento (Brugarolas 1989, Galera 1991).

Con ambas técnicas se saca la conclusión de que las lesiones mamarias benignas son generalmente diploides (Enchev 1987, Christov 1989, Palmer 1988). Sin embargo los tumores



---

mamarios malignos son diploides sólo en una proporción que oscila entre el 29% al 35% según autores, teniendo el resto una dotación anormal de cromosomas (Spyratos 1987, Cornelisse 1987, Christov 1989). Los tumores diploides presentan diferencias de pronóstico con los aneuploides (Dowle 1987, Kallioniemi 1988), es menos frecuente la afectación ganglionar y con más frecuencia son receptor de estrógenos positivos (Berryman 1987) (características que presentan un mejor pronóstico). En contraposición las pacientes con tumores malignos aneuploides son con más frecuencia receptor de estrógenos negativos (Owainati 1988, Cornelisse 1987), tienen un alto índice mitótico (Uyterlinde 1987) y una alta fracción S con respecto a los diploides (Dressler 1988, Kallioniemi 1988, Muss 1989), características relacionadas con peor pronóstico.

Se da también valor pronóstico a la proporción de células de cáncer de mama en las que se puede identificar CORPÚSCULOS CROMATÍNICOS SEXUALES (Rajeswari 1977), encontrando Savino y Koss (1971) que mujeres con un 20% o más de células cancerosas con estos corpúsculos presentan mayor supervivencia.

Con respecto a la MORFOMETRÍA NUCLEO-CITOPLASMÁTICA la bibliografía recoge las siguientes mediciones en las células cancerosas mamarias con valor pronóstico: diámetro nuclear (Zajdela 1979, van Bogaert 1980) área nucleolar (Baak 1985), área nuclear (Maehle 1982, Pienta 1991) y perímetro, ejes, tamaño



---

y factor de forma (Baak 1982, Linden 1986) y tamaño de los grupos celulares malignos (Walgren 1976).

En la literatura van apareciendo otros métodos de cuantificación de la actividad celular como la cuantificación de PROTEÍNAS NUCLEARES ASOCIADAS A LA ACTIVIDAD CELULAR tales como: Ki67, P145 nucleolina y proteínas argirófilas asociadas a las regiones organizadoras del nucleolo (NOR-AP), con ésta última técnica se ha encontrado diferencias notables en el número de nucleolos entre cánceres de bajo y alto grado (Crocker 1987, Hall 1988) o entre lesiones mamarias benignas y malignas (Smith 1988).

Otros factores cito-histológicos, como los niveles de Catepsina C (Tandon 1990) y de marcadores tumorales, fundamentalmente fosfohexosaisomerasa (PHI) y carcinoembrionario (CEA), también orientan el pronóstico (Gonzalez-Merlo 1988).

### **I-3-3-C Factores endocrinos: Receptores hormonales**

La falta de receptores estrogénicos puede indicar una alta malignidad y su recurrencia es más frecuente en los tejidos blandos, mientras que los receptores positivos la tienen en el tejido óseo (Clark 1987).



---

inversamente proporcional a su contenido en receptores de estrógenos; tumores receptor negativo (RE-negativo) recidivan antes y, en general, tienen peor pronóstico (aproximadamente el 50% de los cánceres de mama son negativos). Se conoce una relación inversa entre el contenido de receptores de estrógenos y receptores para el Factor de Crecimiento Epidérmico (REGF). Parece que la activación del REGF estimula la liberación de una Uroquinasa cuyo poder fibrinolítico aumenta la invasividad tumoral (Brugarolas 1989). Los RE-negativo también se los ha relacionado con núcleos más grandes y alto índice de DNA (Larsimont 1989).

### **I-3-3-D Factores genéticos: Expresión de oncogenes**

La existencia de amplificación de determinados oncogenes en tumores malignos de mama se ha relacionado con el pronóstico (también con la afectación de los ganglios linfáticos por el tumor y los receptores hormonales). Los más frecuentes son la amplificación del oncogén Her-2/neu y el c-myc que se presentan en mujeres con cáncer de mama con una frecuencia entre el 20 al 30% el primero y el 6 al 15% el segundo (Seshadri 1989, Guerin 1989). Parece que el producto del gen HER-2/neu codifica para una proteína quinasa (Kolata 1987), enzima que fosforila tirosinas. La fosforilación de tirosinas es un mecanismo normal en la regulación del crecimiento celular pero muchos investigadores piensan que ésta función la comparten los enzimas de oncogenes (Bishop 1985),



---

mimetizando así algún mecanismo de control celular mediante la fosforilación de una proteína que normalmente se produce por la acción de un receptor (Hunter 1985).

#### **I-4 METODOS DE ESTUDIO CONVENCIONALES DE LA PATOLOGIA MAMARIA**

Existe una gran cantidad de métodos de estudio para el diagnóstico de la patología mamaria pero muchos de ellos como la Tomografía Computarizada, la Termografía, la Resonancia Magnética Nuclear o la Dactanografía, tienen poco valor diagnóstico por sí mismos, son demasiado caros para los resultados que ofrecen o están aún en fase experimental.

Los métodos más utilizados son: LA EXPLORACIÓN CLÍNICA, LA MAMOGRAFÍA, LA ECOGRAFÍA Y LA CITOLOGÍA.

García-Valdecasas (1988) y Marcos (1986) en un estudio comparativo que incluía la exploración clínica, la Mamografía y la Citología concluyen que, como prueba diagnóstica aislada, es la Citología la que consigue diagnosticar más número de cánceres de mama, divergiendo uno y otro autor en la prueba que se sitúa en segundo lugar. Ambos coinciden en que la combinación de algunos o de todos los métodos eleva el valor diagnóstico.



---

#### I-4-1 EXPLORACION CLINICA

Se trata de la exploración manual de la mama por el médico especialista. Es el método más rentable y sencillo. Con ella se determina la existencia de la tumoración, su tamaño, movilidad, dureza y existencia de secreción mamaria. El menor tamaño de un tumor que puede ser percibido suele medir aproximadamente 5mm. de diámetro, sin embargo en la mayoría de los casos sobrepasa los 0,5 cm. Con éste tamaño los tumores malignos presentan ya fibrosis peritumoral que se irradia a los tejidos vecinos y a la piel modificando el aspecto de ésta. Esta adherencia tumoral hace que el tumor maligno no tenga unos contornos precisos lo cual lo distingue de un tumor benigno (Puente 1976).

El porcentaje de diagnósticos correctos con éste método varía desde el 61% al 88% según autores (Egan 1976, Garrido-Teruel, García-Valdecasas 1988).

#### I-4-2 MAMOGRAFIA

Es la impresión de la mama en una película de alto contraste. Para ello utiliza rayos X de baja energía que pone de manifiesto las diferencias de densidad que existe en la mama (Doust 1986). Hasta hace poco tenían riesgo por utilizar radiaciones ionizantes y, por tanto, con posibles efectos cancerígenos (Carlile 1981) pero los nuevos aparatos han



---

minimizado estos efectos y han aumentado el valor diagnóstico, principalmente en la detección de tumores de pequeño tamaño no detectados aún clínicamente.

En la mamografía se observan dos tipos de lesiones: las circunscritas y las estrelladas, las primeras orientan hacia un diagnóstico de benignidad y las estrelladas hacia la malignidad (Gonzalez-Merlo 1988).

La exactitud diagnóstica depende de la edad de la paciente y de la etapa de la enfermedad, siendo menos eficaz en las mujeres premenopáusicas, donde la glándula es más densa, que en las menopáusicas (Strax 1977).

#### **I-4-3 ECOGRAFIA**

Es la imagen que se produce por ultrasonidos. Estos son una vibración mecánica de alta frecuencia de igual naturaleza física que el sonido audible aunque de mayor frecuencia (alrededor de los 2,25 millones de ciclos por segundo). A diferencia de las radiaciones electromagnéticas, como los rayos X utilizados en las Mamografías, el nivel de energía utilizado en la Ecografía no ioniza ni produce radicales libres en los tejidos y, por tanto, es un medio diagnóstico bastante inocuo pero con un valor diagnóstico inferior a la mamografía y la citología por punción aspiración (Doust 1986).



---

#### I-4-4 CITOLOGIA POR PUNCION ASPIRACION

Esta técnica comenzó a utilizarla Greig y Gray en 1904 en ganglios linfáticos, empezó a difundirse gracias a los trabajos de Martin y Ellis (1934) en el Memorial Hospital de New York y a ser más accesible su conocimiento cuando en 1946 aparecieron los primeros criterios citológicos de malignidad de Jackson y Severance (1946).

En una tabla de rendimientos de diversos autores recogida por Jiménez-Ayala (1988) la Sensibilidad del método oscila entre el 90% y el 97% y la Especificidad entre el 94% y el 99.7%.

La técnica es fácil y económica, no es necesario poner anestesia ya que la punción es poco dolorosa y se realiza con una aguja similar a la utilizada para la inyección intramuscular conectada a una jeringa de 10 o 20 cm. Debe hacerse de un sólo golpe y rápidamente, realizando el vacío suavemente para extraer el material hacia la jeringa, dejando de succionar al retirarla para evitar la siembra de células cancerosas a lo largo del recorrido. A continuación el material se extiende sobre un porta y se fija con el fijador apropiado según el tipo de tinción que se vaya a emplear.

El extendido recuerda la arquitectura histológica



---

de donde proviene. Los grupos de células ductales benignos se disponen generalmente en monocapa conservando básicamente la estructura que tienen en el epitelio, la succión y extensión contribuye a su fragmentación, los grupos de células malignos, generalmente más pequeños por la pérdida de cohesión debido a la malignización, aparecen con diferentes grados de desorden como corresponde a un crecimiento anárquico.

#### **I-4-4-1 Criterios de malignidad citológico**

Existen unos criterios generales de malignidad, pero cada tipo de cáncer tiene unas características específicas. Los criterios no son absolutos sino que es la unión de diversos factores aprendidos con la experiencia propia y de otros profesionales lo que determina un conocimiento empírico global de malignidad. Estos "INDICADORES" van a orientar sobre la naturaleza del proceso y van a ser barajados por el conocimiento del experto al dar el diagnóstico. La valoración debe hacerse sobre material bien conservado y teñido, teniendo en cuenta que determinados procesos de degeneración y muerte celular pueden simular malignidad. Los criterios más objetivos los van a proporcionar el núcleo más que el citoplasma (Novak 1970).



---

#### I-4-4-1-A Criterios Nucleares

##### -TAMAÑO Y FORMA :

Son factores destacados en los frotis mamarios por Wilson 1978 y Shabaik 1993. Orientan a la malignidad los núcleos agrandados (Takahashi 1973, Gadennne 1985, Battacharjee 1985, Witteking 1987, Detweiler 1988), con diferencias de tamaño en un mismo grupo (anisocariosis), efecto producido por el crecimiento anárquico de las células. Sin embargo el pleomorfismo celular también puede darse en algunas lesiones mamarias benignas y ser causa de falsos positivos (Wilson 1978).

##### - HIPERCROMASIA:

Es usual pero no es la regla ya que la degeneración nuclear es causa frecuente de hipercromasia. La falta de nitidez del perfil nuclear orientará hacia una hipercromasia por degeneración. La diferencia de tinción dentro de una misma muestra puede también ser causa de error (Novak 1979).

##### -DISTRIBUCIÓN CROMATÍNICA

Los núcleos malignos presentan una cromatina grumosa y desordenada, una distribución regular orienta hacia la benignidad. La aglomeración cromatínica que origina en el núcleo zonas claras y oscuras y la angulación brusca de éstos orientan hacia la malignidad (Novak 1970, Takahashi 1973). Siendo la irregularidad de la cromatina uno de los signos más importantes de malignidad en la citología mamaria (Wilson 1978,



---

Gadenne 1985, Detweiler 1988).

-LA IRREGULARIDAD DE LA MEMBRANA NUCLEAR

Membrana gruesa y/o con irregularidad en el grosor generalmente con acúmulos de cromatina adherida a ella de forma variable (Novak 1970). Este proceso es también frecuente en la degeneración celular (Takahashi 1973), en éste caso otros signos de degeneración deben acompañarlo .

- NUCLEOLOS

La observación de las características nucleolares solo es orientativa (Wilson 1978). Gran diferencia en el número de nucleolos en los núcleos de un mismo grupo, aumento del tamaño de éstos e irregularidad en el contorno de los mismos sólo servirán de apoyo a otros factores de malignidad más objetivos, ya que procesos regenerativos también pueden presentarlos y, de hecho, Wilson (1978) y Azúa (1987) señalan la sobrevaloración de los nucleolos como causa de falsos positivos ya que su agrandamiento puede darse en procesos benignos mamarios por aumento de la actividad.

-DESORDEN NUCLEAR

La alteración de la forma de los núcleos por la proximidad de los núcleos vecinos y la irregularidad en su disposición es consecuencia del crecimiento rápido y desordenado (Novak 1970) y es de especial orientación en la citología mamaria (Wilson 1978, Dziura 1979, Wolberg 1987, 1989).



---

#### I-4-4-1-B Criterios Citoplasmáticos

La cantidad de citoplasma tiende a ser menor en las células malignas, generalmente a causa del aumento de tamaño del núcleo, aunque también puede deberse a una disminución del citoplasma (Novak 1970, Bentura 1988, Jimenez-Ayala 1988). Se estima relacionando el área del núcleo con el área del citoplasma (Boon 1982, García-Valdecasas 1988) y es más valorable en células de tamaño mediano y grande, teniendo poco valor en células pequeñas y con escaso citoplasma (Novak 1970). Esta relación permanece dentro de unos límites constantes en las células normales (TaKahashi 1973).

Otras características más generales del extendido han sido valoradas para el diagnóstico citológico de malignidad en la mama: la gran celularidad, pérdida de cohesión de las células en los bordes de los grupos, la gran cantidad de células epiteliales sueltas, la ausencia de células mioepiteliales o apocrinas, la presencia de células inflamatorias (Wolberg 1987 y 1989, Jimenez Ayala 1988, Hutchinson 1991, Shabaik 1993). (Figura 5).



---

Figura 5a: EXTENSIÓN CITOLÓGICA DE LESIÓN BENIGNA MAMARIA OBTENIDA POR PUNCIÓN-ASPIRACIÓN

(imagen tomada a 100 aumentos)

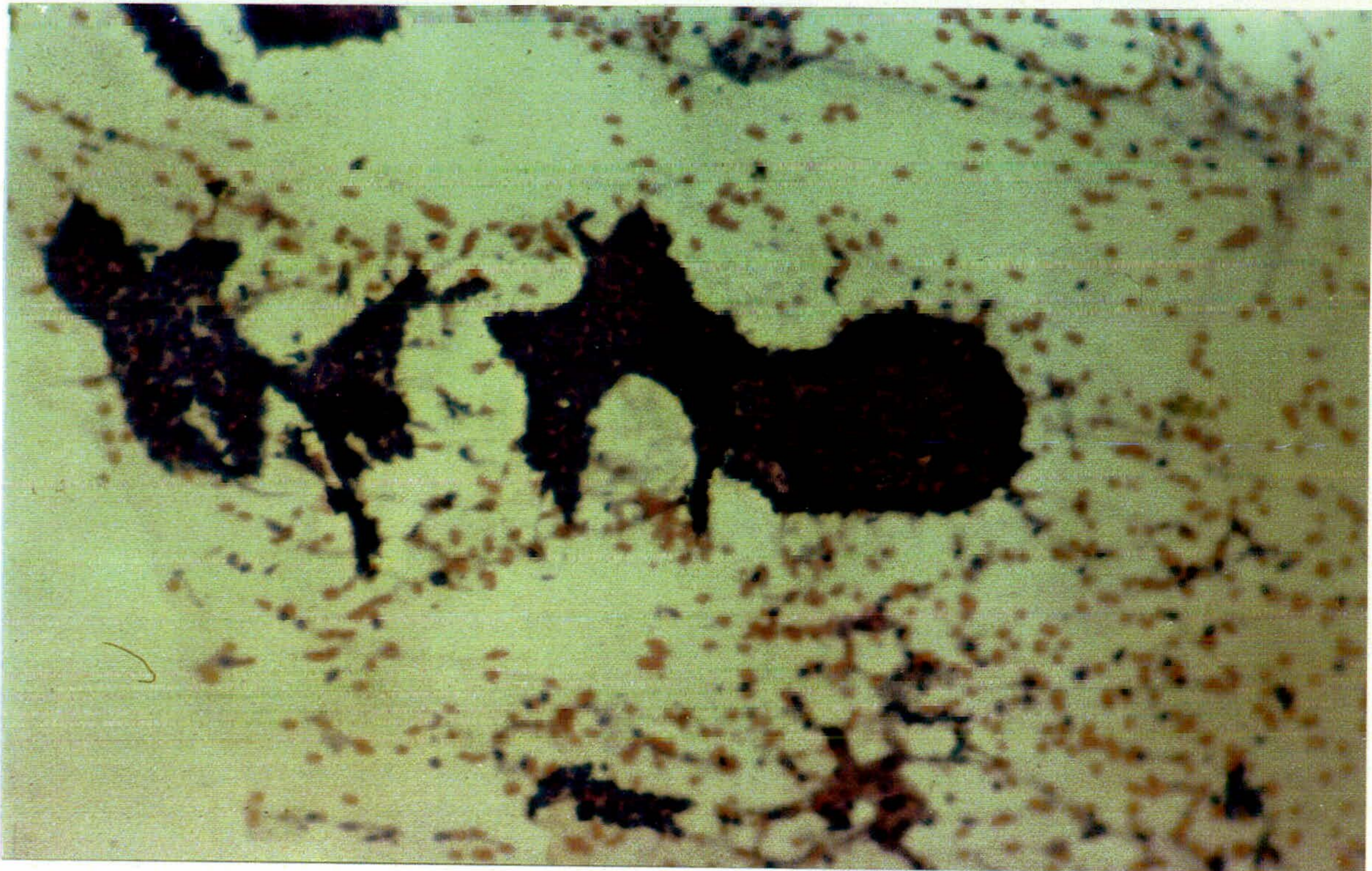
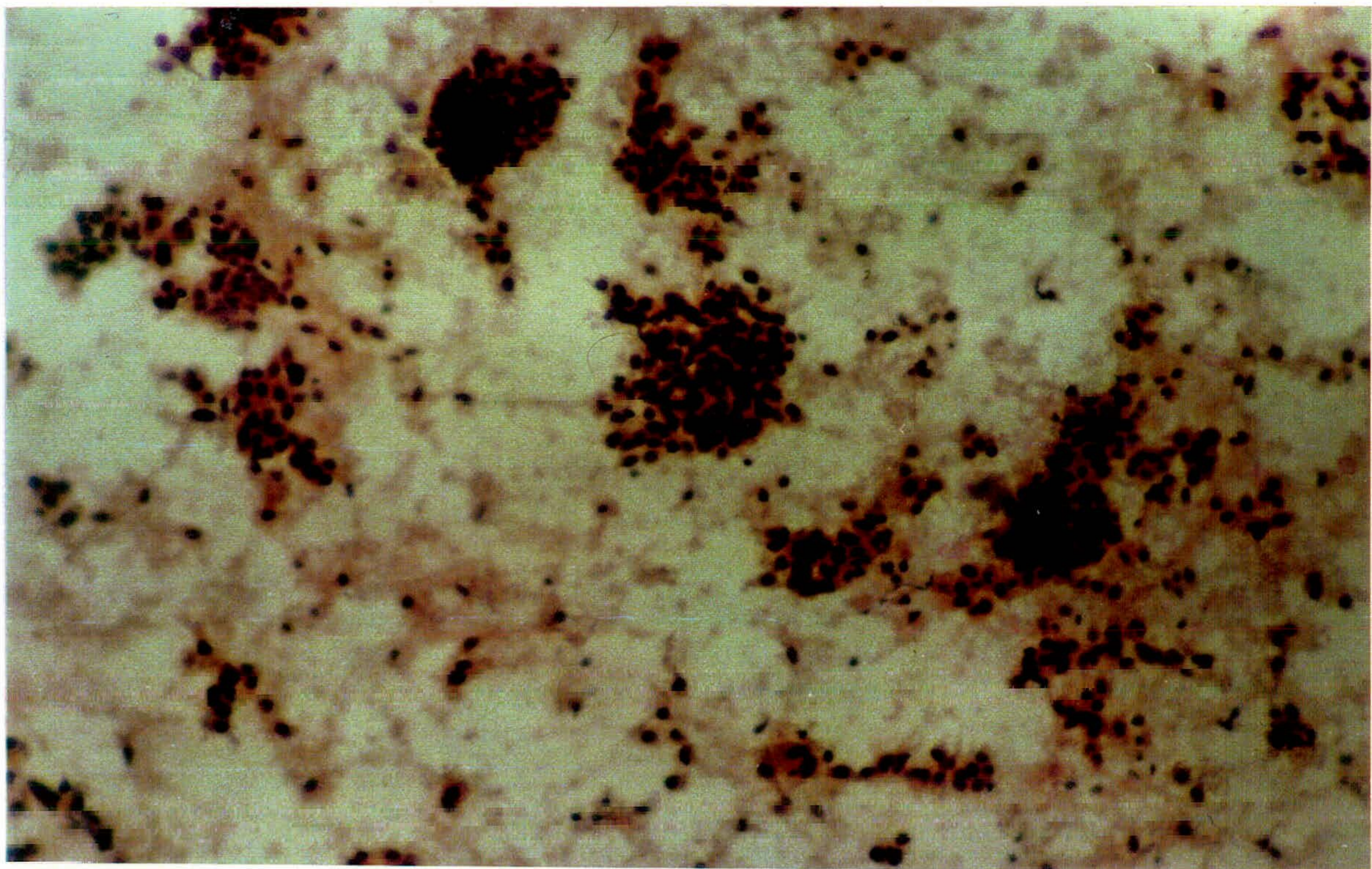


Figura 5b: EXTENSIÓN CITOLÓGICA DE LESIÓN MALIGNA MAMARIA OBTENIDA POR PUNCIÓN-ASPIRACIÓN

(imagen tomada a 100 aumentos)





---

## **I-5 METODOS DE ESTUDIO NO CONVENCIONALES DE LA PATOLOGIA MAMARIA: METODOS CUANTIMETRICOS**

Con ellos se obtiene datos numéricos de los estudios anatomopatológicos, tendiéndose a restringirlo al nivel microscópico. Brugal (1986) lo define como "Patología cuantitativa".

Los métodos de valoración cuantitativa se dividen en (Ferrer-Roca 1986):

- 1 - *Fotometría o Densitometría*
- 2 - *Morfometría*
- 3 - *Análisis textural*
- 4 - *Análisis Computarizado de la Imagen*  
(incluye todos los anteriores)

### **I-5-1 METODOS DE ESTUDIO FOTOMETRICOS O DENSITOMETRICOS**

Su finalidad principalmente es la cuantificación del DNA y se puede llevar a cabo por dos métodos: LA MICROESPECTROFOTOMETRÍA Y LA CITOMETRÍA DE FLUJO.

#### **I-5-1-A Microespetrometría o Citofotometría:**

Se inició por Casperson y col. en la década de los 30 y se basa en el revelado del elemento bioquímico a medir



---

mediante una técnica específica y posterior cuantificación de la sustancia. Para ello es imprescindible que la reacción sea estequiométrica y así poder aplicar la Ley de Beer-Lambert que se basa en el principio de que "la extinción de la luz es directamente proporcional a la concentración de la coloración en una sustancia determinada":

$$\text{Log } I_0/I = K d c$$

Siendo  $I_0$  = luz incidente,  $I$  = Intensidad de la luz que atraviesa el objeto,  $K$  = Coeficiente de Absorción de la radiación electromagnética,  $d$  = espesor de la cubeta (o espesor entre porta y cubreobjetos),  $c$  = concentración molar

Para la cuantificación de la sustancia se elige una luz monocromática de una determinada longitud de onda, que dependerá del espectro de absorción de la sustancia a valorar y que para la cuantificación del DNA, utilizando la reacción de Feulgen, se encuentra entre 550-575 nm. (Bentura 1988).

Existen muchas reacciones específicas para las sustancias biológicas: La reacción de PAS para polisacáridos con grupos 1-2 glicósidos, la reacción de Millon para la Tirosina, la reacción de Feulgen para el ADN...etc. Otros autores sin embargo han utilizado técnicas no específicas, como tinción con hematoxilina, para análisis fotométrico (Pérez-Iraola 1989, Esquivias 1994).



---

### 1-5-1-B Citometría de Flujo:

Se basa en la cuantificación de la fluorescencia que emite una suspensión de células que fluye por un tubo capilar al pasar por un haz de luz (generalmente producido por un láser) y cuya sustancia a valorar se encuentra unida a un fluorocromo. Esta técnica también puede ser utilizada para la identificación de las células en la fase S del ciclo celular y por tanto poder determinar la capacidad proliferativa del tumor.

### I-5-2 METODOS DE ESTUDIO MORFOMETRICOS

Fué entre los años 1920 a 1930 cuando comenzaron los citólogos a señalar cambios morfológicos o metabólicos en los casos de malignidad (Bentura 1988). Los métodos para medir éstos cambios morfológicos pueden ser muy sencillos: retículas oculares o regla y microproyector, habiéndose aplicado en ciencias diversas (Chalkley 1949, Weibel 1963, Hally 1964, Delesse 1984). Estos métodos tradicionales de medida han sido progresivamente sustituidos por métodos de análisis semiautomáticos o automáticos que requieren el uso de ordenadores con hardware y software específicos y donde es condición necesaria la digitalización de las imágenes. Los métodos semiautomáticos cuantimétricos realizan la extracción o segmentación de la zona de interés mediante el dibujo manual con bolígrafo magnético, ratón, tableta digitalizadora



---

o peritérico similar, los métodos automáticos requieren algoritmos complicados de segmentación (Moragas 1985, Rivera-Pomar 1985, Marchevsky 1987, Bentura 1988).

La Morfometría abarca tres métodos (Ferrer Roca 1986):

1-PLANIMETRÍA: Cuantificación de estructuras en dos planos.

2-ESTEREOLOGÍA: Intiere medidas tridimensionales a partir de parámetros bidimensionales.

3-CONTAJE Y LA RECONSTRUCCIÓN SERIADA

La nomenclatura propia abarca términos como: área, fracción del área total y del volumen total, número de elementos discretos por unidad de área y por unidad de volumen, densidad de superficie, diámetro, factor de forma, factor de estericidad, factor de excentricidad, elípticidad, factor de regularidad...

### I-5-3 ANALISIS TEXTURAL

Moragas (1991) define la textura en biomedicina como "la peculiar organización o disposición de ciertos componentes hísticos o citológicos (elementos estructurantes) de la imagen microscópica" y el análisis textural como "la cuantificación numérica de una estructura que permita determinar en qué medida dicha textura se aparta del patrón normal".



---

Las técnicas de análisis textural fueron iniciadas por Wied (1968) para la identificación de células, en un programa conjunto de los departamentos de Microbiología de la Universidad de Arizona, el de Ginecología y Obstetricia de la Universidad de Chicago y el Instituto de Patología de las Fuerzas Armadas de Washington. Posteriormente se ha utilizado para el estudio de alteraciones celulares malignas y premalignas (Rosenthal 1982, Martín 1982, de Torres 1988, Hutchinson 1991, Wolberg 1993) y otras alteraciones estructurales orgánicas (Moragas 1983, de Torres 1986). La mayoría del análisis textural que se realiza hoy día se lleva a cabo mediante las técnicas de análisis de imagen que se expondran más adelante.

Las aproximaciones más importantes al análisis textural son (Ferrer-Roca 1986):

- ANÁLISIS MARKOVIANO: Calculan matrices de probabilidad (o de coocurrencias) de transición de unos niveles de intensidad a otros, entre píxeles cercanos.
- ANÁLISIS DE GRADIENTES: Se calculan primeras derivadas (gradientes) y segundas derivadas (laplacianas). Los parámetros texturales serán la media, varianza, sesgo, curtosis... de gradientes y laplacianas.
- ANÁLISIS GRANULOMÉTRICO: Características de los gránulos como tamaño, forma, intensidades...
- ANÁLISIS DE FOURIER: Descomposición de una función en un conjunto de sinusoides.



---

Como vemos el análisis de texturas es muy complejo, ya que, a pesar de que intuitivamente parece simple, resulta bastante complicado definir diferentes tipos de texturas en base a cualidades expresables analíticamente. Así se habla normalmente de propiedades tales COMO UNIFORMIDAD, DENSIDAD, RUGOSIDAD, REGULARIDAD, DIRECCIONALIDAD...ETC., lo que da una idea de la complicación que supone establecer una definición operacional de textura que asuma todo eso. Tal ausencia de ideas universalmente aceptadas se ha reflejado en el gran número de métodos para el análisis de texturas que han proliferado en la literatura (Haralick 1979, Connors 1980, Davis 1980, Moragas 1985 y 1991). Aunque sí parece aceptado por todos que, entre las distintas aproximaciones, dos son las interesantes en biomedicina: la estadística y la estructural. En la primera categoría la textura se representa como un conjunto de medidas en forma de *vector de rasgos*, tales COMO: LAS BASADAS EN RECORRIDOS DE NIVELES DE GRISES, DIFERENCIAS ENTRE LOS MISMOS O LAS BASADAS EN MATRICES DE COOCURRENCIAS. Todas éstas medidas se suelen deducir de forma empírica y la textura original no puede reproducirse a partir de los rasgos. En la categoría estructural las regiones de textura se ven como patrones 2-D que constan de un conjunto de *primitivas* (patrones que se repiten a lo largo de la imagen y que forman la textura) con una determinada relación entre sí (los elementos o *textel* presentan una regularidad que puede definirse por un patrón de comportamiento al contrario que en el modelo estadístico). (Haralick 1979).

---



---

#### I-5-4 ANALISIS COMPUTERIZADO DE LA IMAGEN

El Análisis de imagen incluye metodología para la densitometría, la morfometría y el análisis textural.

##### I-5-4-1 Aplicación a la morfología celular

Esta tecnología aplicada al ámbito de la Biología y Medicina comenzó entre la década de los años 50 y 60. En 1954 Mendelsohn y col. construyeron el CYDAC definiendo patrones de identificación celular. Un proyecto para la construcción de un citoanalizador del National Cancer Institute fué abandonado entre 1955 y 1961 (Wied 1968) porque requería células aisladas y las variables manejadas no fueron efectivas para el diagnóstico. Posteriormente se desarrolló la máquina automática de reconocimiento de leucocitos CELLSCAN de Preston (1961) y el FIDAC (Ledley 1966). Estos tres sistemas utilizan el análisis métrico del contorno de los objetos para su reconocimiento. Sin embargo los trabajos de Preston (1961) y posteriormente Prewitt y Mendelsohn (1966) eran incompletos pues solo podían analizarse unos cuantos cientos de células. A partir de este momento se desarrollaron numerosos analizadores, uno de los más completos fué el microespectrofotómetro UMSP-I de Carl Zeiss que incorporó un mecanismo de alta precisión basado en la utilización de la luz ultravioleta, constituyendo éste el núcleo del TICAS (Taxonomic Intracellular Analytic System) (Wied 1968). Estas



---

técnicas, que utilizaban televisión acoplada al microscopio, fueron rápidamente reemplazadas en Hematología por el contaje de células a través de un fluido, basados en el principio de Coulter (1956).

Actualmente existe en el mercado diversos software de aplicación en biomedicina como el TICAS (Wied 1968), el CYBEST (Watanabe 1974), el TEXCAN (Ferrer-Roca 1990)... Este último, utilizando técnicas de Morfología Matemática, realiza automáticamente la gradación tumoral de Carcinomas tiroideos y de mama.

#### I-5-4-2 Metodología

La metodología en el análisis computerizado de la imagen requiere inicialmente la digitalización de ésta, pasando de ser una función continua a una función discreta. A la discretización espacial se la denomina muestreo sucesión de un número fijo de puntos de luz llamados píxeles y a la discretización en amplitud: cuantificación del nivel de gris, que es la asignación a cada píxel de un valor numérico dependiendo de la intensidad luminosa que genera cada punto. Por tanto una imagen digital es una representación bidimensional de un determinado número de puntos, cada uno de ellos definido por una función  $f(x,y)$  que determina un valor numérico (en función de la intensidad de luz), asignado a una posición en un plano definido por un





---

sistema de coordenadas.

Los componentes principales de un Analizador de Imagen son:

- 1 - UN ORDENADOR CON HARDWARE Y SOFTWARE ESPECIALIZADO
- 2 - UN SISTEMA DE CAPTACIÓN DE LA IMAGEN (UNA VIDEO CÁMARA POR EJEMPLO).
- 3 - UN DISPOSITIVO PARA LA VISUALIZACIÓN DE LA IMAGEN (UN VIDEO MONITOR POR EJEMPLO). (Figura 6).

Figura 6. Analizador de imagen utilizado en éste trabajo





---

El hardware consta fundamentalmente de un Convertidor analógico digital (ADC) que convierte señales analógicas (voltajes) en digitales (valores numéricos). El sistema debe tener un sistema de almacenamiento (buffer) que es una memoria adicional conectada a la salida del ADC y que puede tomar, al menos, una imagen digitalizada a la vez.

Las etapas fundamentales en el análisis de imagen son:

- A- LA CAPTACIÓN DE LA IMAGEN
- B- EL PROCESAMIENTO DE LA IMAGEN
- C- EL ANÁLISIS DE LA IMAGEN

#### a-Captación de la imagen

Generalmente se realiza con sensores para la captación fotolumínica como los microdensitómetros (fotómetros, reflectómetros...) o cámaras de televisión. Estas últimas con la ventaja de realizar conversión analógico-digital de la imagen y no introducir distorsión geométrica en la resolución lineal (Ferrer-Roca 1990). En particular las cámaras CCD (Charge Couple Device) tienen una relación lineal entre la luminosidad media y la verdadera en cada pixel, por ello son muy adecuadas en la captación de la imagen. La cámara de video emite en un número de líneas paralelas con una sucesión de diferentes niveles de gris, el Convertidor Analógico Digital divide cada línea en un número fijo de



---

puntos de luz o píxels y asigna un valor numérico a cada nivel de gris dependiendo del voltaje que genera cada punto (Marchevsky 1987).

### b-Procesamiento de la imagen

Es la realización de algún tipo de transformación sobre la imagen con objeto de hacer mejor el análisis. Por su utilización en éste trabajo se hará mención a funciones de REALCE, FILTRADO Y SEGMENTACIÓN.

El realce es la modificación de la imagen para obtener otra más adecuada al análisis del experto o del ordenador. Abarca operaciones de: alteración del contraste, modificación del histograma, realce de contornos..etc., en los que muchas veces se utilizan técnicas de filtrado.

El filtrado, en nuestro contexto, está relacionado con la detección de fronteras en la imagen con la finalidad de separar distintos objetos presentes. Son muchas las técnicas desarrolladas para este fin: detectores basados en el gradiente, segundas derivadas, ajuste de superficies o en técnicas multiresolución.

Con la segmentación se subdivide la imagen según los niveles de gris que presente obteniéndose regiones o contornos. Para ello la binarización suele ser la función más



---

empleada. Con ella, a partir de un determinado umbral, la imagen se representa con un nivel y por debajo de dicho umbral de otro. El procedimiento puede ser automático o interactivo. Este último es el más habitual, aunque el trazo manual es desigual y el procedimiento muy lento por lo que su interés está limitado a casos en que la imagen tenga objetos altamente contrastados. El procedimiento automático utiliza técnicas basadas en las propiedades de los niveles de gris: *la similaridad y la discontinuidad*. En cualquier caso esta umbralización, sea manual o automática, no suele ser suficiente por sí misma para determinar las diferentes regiones de interés presentes en la imagen.

#### c-El Análisis de la imagen

Es la interpretación de la imagen y comprensión de sus distintas partes. El análisis será diferente según se desee analizar la forma, la distribución de intensidades o la textura.

Con el análisis de la forma se obtendrán los parámetros morfométricos ya mencionados: ÁREA, PERÍMETRO, FACTORES DE FORMA...etc. Con el análisis textural lo que se obtiene es EL HISTOGRAMA DE NIVELES DE GRIS CON EL RANGO, MODA, MEDIANA, ESTADÍSTICOS DE ORDEN 2 COMO DIFERENCIAS DE NIVELES DE GRIS, COOCURRENCIAS DE NIVELES DE GRIS..



---

Finalmente en este apartado dedicado al análisis merece un comentario particular las técnicas basadas en la MORFOLOGÍA MATEMÁTICA (Matheron 1967, Serra 1982, Moragas 1983, Cornélisse 1985, De Torres 1986, Giménez Mas 1991) que intuitivamente basan el procesamiento y análisis de la imagen en el concepto de forma mediante la modelización de los objetos presentes en la imagen. La mayoría de los algoritmos en morfología trabajan fundamentalmente con dos transformaciones base: la erosión y la dilatación, mediante un ELEMENTO ESTRUCTURAL, obteniéndose otra imagen transformada (Serra 1982). Este elemento estructural modifica la imagen en virtud del grid elegido siendo el rectangular el más usual. Con la combinación de éstas dos transformaciones se obtienen otras más complejas como la apertura (erosión seguida de dilatación), la clausura (dilatación seguida de erosión) o el engrosamiento y adelgazamiento. En general son operaciones destructivas porque la imagen inicial no puede recuperarse a partir de la imagen transformada. Por ejemplo: la apertura alisa contornos, rompe istmos y elimina protusiones finas y la clausura elimina agujeros y rellena saltos en el contorno. Todas éstas transformaciones se pueden hacer tanto con la imagen en binario como en niveles de gris. Así mismo existen filtros morfológicos cuya característica esencial es la no linealidad, frente a los que normalmente se utilizan que suelen ser lineales.



---

## **I-6 METODOS DE CLASIFICACION DE DATOS**

La construcción de un procedimiento de clasificación a partir de un conjunto de datos para los cuales se conocen clases verdaderas se ha denominado históricamente RECONOCIMIENTO DE PATRONES, DISCRIMINACIÓN o APRENDIZAJE SUPERVISADO, con objeto de distinguirlo del APRENDIZAJE NO SUPERVISADO en el que las clases se infieren a partir de los datos.

### **I-6-1 TAXONOMIA DE LOS METODOS DE CLASIFICACION**

Se pueden diferenciar históricamente tres bloques de métodos: ESTADÍSTICOS, DE APRENDIZAJE Y REDES NEURONALES.

#### **I-6-1-A Aproximaciones Estadísticas**

Existen dos tendencias: la clásica, concentrada en la discriminación lineal y sus variantes y la moderna, que explora clases de modelos más flexibles, la mayoría de los cuales intentan proporcionar un estimador de la probabilidad conjunta de los rasgos dentro de cada clase. Por la importancia y su uso en datos biomédicos se describirán brevemente las características de un método clásico: el Análisis Discriminante.



---

## Análisis Discriminante

La idea es encontrar la combinación de variables que mejor predican la categoría o el grupo al que cada caso pertenece. La combinación de tales variables predictoras se denomina función de clasificación y será usada para asignar los casos nuevos en el grupo al que pertenecen. El primer problema es la selección de variables que, a priori, se piense que puedan tener importancia en el proceso de clasificación, y el segundo es la forma operativa del análisis. En general se suelen calcular las funciones de clasificación basándose en un subconjunto de los datos y se usan éstas funciones para clasificar los datos restantes. Para ello se subdividen los casos de cada grupo en dos subgrupos, el primero se usa para estimar el clasificador y el segundo para validarlo. Observando la proporción de casos del segundo subgrupo bien clasificados tendremos una medida empírica del éxito de la discriminación. Analíticamente, para cada objeto (célula, imagen, grupo...) se determina un vector de rasgos  $x = \{x_1, \dots, x_p\}$  de forma que las funciones discriminantes se escriben como:

$$f_i(x) = w_{i0} + \sum_j w_{ij} x_j$$

Siendo  $i=1, \dots, nm\_clases$ ,  $w_i = \{w_{i1}, \dots, w_{ip}\}$  un vector de pesos y  $w_{i0}$  una constante, de manera que el clasificador asignará un vector de rasgos  $x$  a la clase  $w_i$  si  $f_i(x) > f_j(x)$  para todo  $j$  diferente de  $i$ . Tal clasificador divide el espacio de rasgos en  $(nm\_clases)$  regiones de decisión, con  $f_i(x)$  el mejor discriminante si  $x$  está en la región  $i$ -ésima.



---

## **I-6-1-B Aprendizaje**

El aprendizaje en éste contexto se relaciona con procedimientos informáticos basados en operaciones lógicas o binarias que aprenden una tarea a partir de una serie de ejemplos. De todos los métodos existentes merecen especial atención los los árboles de decisión, por su uso creciente en datos biomédicos (Mui 1980, Landeweerd 1981, Lin 1983, Marchevsky 1986, Wolberg 1987, 1988 y 1989). En ellos la clasificación es el resultado de una secuencia de pasos lógicos. Otras técnicas como los algoritmos genéticos y los procedimientos lógicos inductivos están en desarrollo con la idea de tratar con tipos más generales de datos, incluyendo casos donde el número y tipo de atributos puede variar, o casos con atributos estructurados jerárquicamente. Por su importancia y por la novedad que supone su uso en datos de la naturaleza de los estudiados en esta memoria, se describirá brevemente la metodología de los Árboles de Clasificación.

### Árboles de Clasificación

Los árboles binarios de clasificación proporcionan una forma interesante de análisis de datos en problemas de clasificación (Breiman 1984) y constituyen una potente alternativa a los métodos paramétricos que realizan la misma tarea. La idea es que una vez que se tienen una serie de medidas para un objeto (célula, imagen, grupo...) en forma



---

de vector, se construye una regla de predicción que asigne cada objeto a una clase. En la construcción de tales reglas de predicción hay un principio fundamental a seguir: construir la regla de predicción más exacta y más intuitiva (fácil de interpretar) posible. La causa está en que a menudo en un problema de predicción se miden muchas variables, pero la cuestión fundamental es determinar qué variables medidas dan información significativa y cuales no. En cualquier caso, la ventaja de ésta aproximación sobre la anterior está en que las reglas de predicción no vienen dadas en forma de expresiones algebraicas a veces difíciles de comprender e interpretar, sino que se construyen en forma de árboles binarios de decisión, fáciles de comprender, usar y explicar. La precisión que se alcanza con los árboles de clasificación es similar a la de los métodos de análisis multivariante e incluso da mejores resultados cuanto más complejas y numerosas son las variables que se manejan (Fernández-Valdivia 1991).

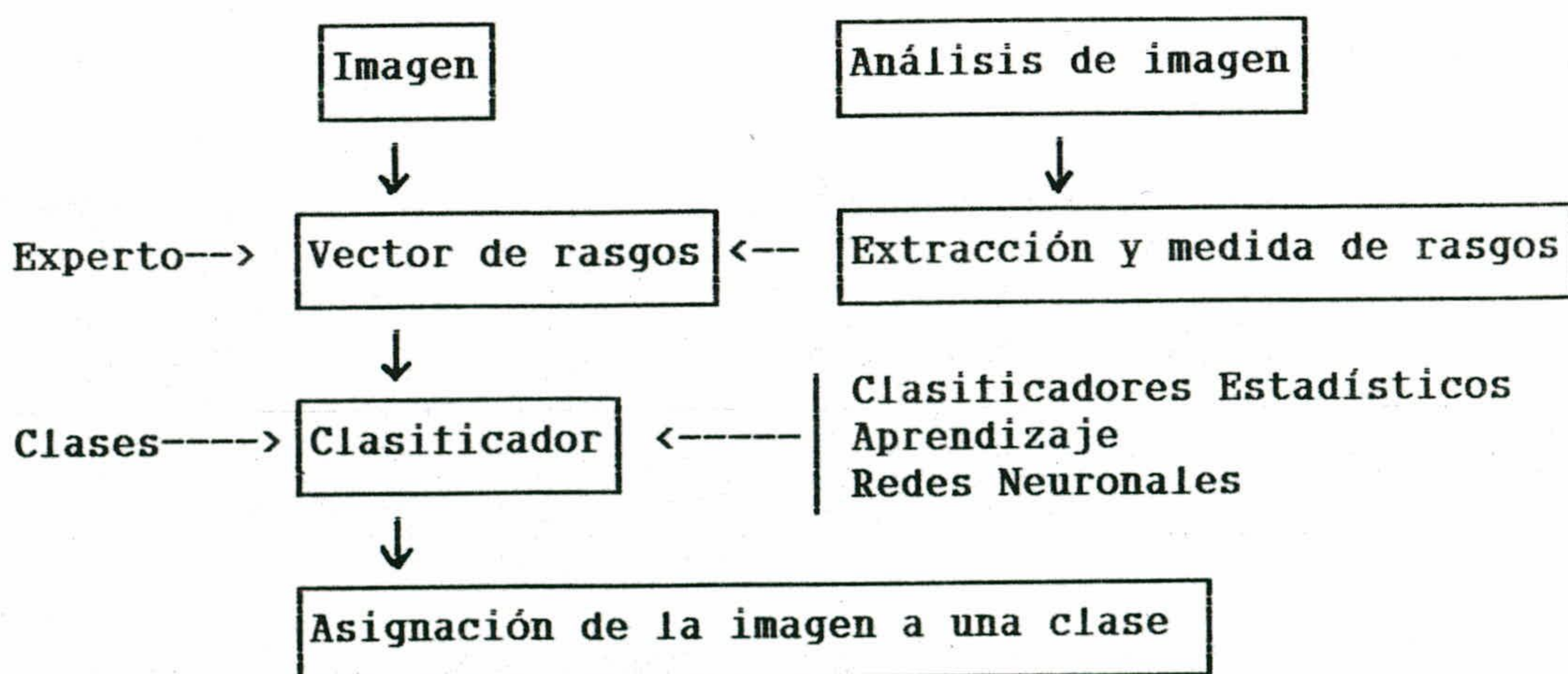
### **I-6-1-C Redes Neuronales**

En términos generales, una red neuronal no es más que una red con nodos interconectados, cada uno de los cuales produce una función no lineal de su entrada. La entrada a un nodo puede venir de otros nodos o directamente desde los datos de entrada. También se suele identificar a algunos nodos con la salida de la red, por tanto todo presenta un



conjunto muy complejo de interdependencias que puede incorporar cualquier grado de no linealidad. En las redes más simples, la salida de un nodo sirve de entrada a otro nodo de forma que se propagan los "mensajes" a través de la red. En redes más complejas es posible que los nodos de salida se conmuten con nodos anteriores de forma que exista retorno ("feedback"). Esta aproximación por redes neuronales (de la que se han desarrollado algunas aplicaciones a biomedicina como por ejemplo Paul (1989) y Dawson (1991)), combina la complejidad de algunas técnicas estadísticas clásicas con el aprendizaje intentando emular el comportamiento humano. El problema que tienen es la dificultad intrínseca de su construcción junto con su elevado costo computacional (López 1989, Fogelman 1988).

#### ESQUEMA GENERAL DE LA CLASIFICACION ORIENTADA A RASGOS





---

## I-7 OBJETIVOS DEL ESTUDIO

Para llegar a construir el clasificador propuesto, es preciso cubrir los siguientes objetivos:

1) Seleccionar extendidos citológicos de PAAF de lesiones malignas y benignas de mama con comprobación histológica.

2) Definir, tras estudio piloto, las variables cuantificables con técnicas de análisis de imagen.

3) Captación de imágenes de campos suficientemente representativos de los extendidos.

4) Aplicación de algoritmos de análisis de imagen para conseguir una segmentación lo más automatizada posible.

5) Desarrollo del clasificador mediante árbol de clasificación (CART) y Análisis Multivariante, con test de significación, para catalogar las lesiones en benignas o malignas a partir de los datos obtenidos por análisis de imagen.



---

**MATERIAL Y METODO**



---

## II MATERIAL Y METODO

En una relación ordenada por orden alfabético de mujeres con cáncer de mama registrados entre 1986 y 1991 en el Hospital Virgen de las Nieves de Granada, se consideraron dos grupos: de la A al H y de ésta a la Z. Se tomo al azar un grupo (300 casos), del que se eligieron aquellos carcinomas mamarios con diagnóstico citológico comprobado por biopsia y diagnóstico mamográfico (requisito éste último para otros aspectos del proyecto en los que ésta tesis estaba integrada). Del número resultante se escogieron los casos cuyas extensiones se encontraban en condiciones aceptables para el análisis citológico, quedando para el objeto de éste trabajo un total de 95 casos, en su mayoría Carcinomas Ductales y el resto lo componía un pequeño grupo (el 8% del total) de carcinomas de diversos tipos (1 Carcinoma Intraquístico, 4 Carcinomas lobulillares, 2 Carcinomas Medulares, 1 Sarcoma Phyllodes). Como grupo control se escogieron 47 casos de lesiones benignas de mama (la mayoría Mastopatías Fibroquísticas y Fibroadenomas).

De cada caso se rellenó un protocolo de datos clínicos, recogidos de la historia de cada paciente, y un protocolo de datos citológicos. Cada caso constaba aproximadamente de 5 extendidos, lo que hacen un total de 735 frotis.



---

Los extendidos citológicos de las punciones realizadas en los tumores mamarios y teñidos con la técnica de Papanicolaou se examinaron con un microscopio Leitz Laborlux-S. Esta tinción tiene un colorante básico: la Hematoxilina que no tiñe los núcleos estequiométricamente ya que no es una tinción específica para éstos.

El Analizador de imágenes constaba de: un microscopio Dialux 20 EB de Leitz con una cámara de video incorporada CCD Sony, dos pantallas de video: una para mostrar la imagen y otra para la interacción con el software, una unidad central de proceso de 8 Mbyte de memoria RAM, una disquetera para discos ópticos de 127 Mbyte de capacidad y el software necesario para hacer el tratamiento y análisis de las imágenes. Concretamente se usó el programa Visilog (Visilog 1988) y el software desarrollado por el grupo de Análisis de Imagenes del Departamento de Ciencias de la Computación de la Universidad de Granada. Las imágenes tenían 256 filas por 256 columnas de pixels y 256 niveles de gris.

Los rasgos que tenían relevancia diagnóstica fueron recogidas en el protocolo I, que se adjunta. Este protocolo recogía rasgos citológicos discriminadores entre lesiones benignas y malignas de mama: variables de los grupos celulares (tamaño, forma, desorden celular), variables del extendido (diátesis tumoral, inflamación, cantidad de células sueltas semejantes a la de los grupos), variables de los



---

nucleos (grado de atipia, diversidad en el tamaño, características de la cromatina, características del contorno nuclear, características de los nucleolos) y características del citoplasma relacionadas con el tamaño nuclear. De éstos rasgos, sólo se estimó cuantificables con análisis de imagen:

- En los grupos grandes:

Forma  
Tamaño del grupo

- En los grupos pequeños:

Relacion área nuclear/ área citoplasmática (AN/AC)  
Desorden de los núcleos integrantes

- En los núcleos:

Area  
Perímetro  
Ejes mayor y menor  
Factores de Forma  
Variables texturales

De los grupos grandes se definieron dos variables: TAMAÑO Y FORMA. De los grupos pequeños se definió LA VARIABLE AN/AC y cuatro variables que cuantificaban el desorden: MEDIA DE LAS DISTANCIAS ENTRE NUCLEOS, SD DE LAS DISTANCIAS, MÁXIMA Y MÍNIMA DISTANCIA. De los nucleos 50 variables, que se agruparon en tres experimentos según sus características, para medir el poder clasificador que tenía cada uno de ellos: el primer experimento lo componían VARIABLES PLANIMÉTRICAS Y LOS CUATRO PRIMEROS MOMENTOS DEL HISTOGRAMA DE NIVELES GRIS, el segundo lo componían VARIABLES TEXTURALES y el tercero englobaba a ambos.



---

El total de las variables fueron:

VARIABLES

1	Area de la célula
2	Perímetro de la célula
3	Feret de las X de la célula
4	Feret de las Y de la célula
5	Circularidad del núcleo
6	Estericidad    "    "
7	Elipticidad    "    "
8	Media de los niveles de grises
9	Varianza       "       "
10	Simetría: Momento de orden 3
11	Momento de orden 4
12	Desviación del área del núcleo
13	"    perímetro del       "
14	"    feret de las X del   "
15	"    "                    Y del "
16	"    circularidad del    "
17	"    estericidad del     "
18	"    elipticidad del     "
19	"    media de los niveles de gris
20	"    varianza            "       "
21	"    Simetría: Momento de orden 3
22	"    Momento de orden 4
23	Máximo del área
24	"    del Perímetro
25	"    de Feret X
26	"    de Feret Y
27	"    de Circularidad
28	"    de Estericidad
29	Media de la textura: entropía
30	"                    "    : energía
31	"                    "    : máximo
32	"                    "    : Entropía (diferencias)
33	"                    "    : energía       "
34	"                    "    : máximo       "
35	"                    "    : entropía (coocurrencias)
36	"                    "    : energía       "
37	"                    "    : máximo       "
38	"                    "    : contraste     "
39	"                    "    : diferencia inversa (coocurrencias)
40	Desviación Media de la textura: entropía
41	"                    "    "    : energía
42	"                    "    "    : máximo
43	"                    "    "    : Entropía (diferencias)
44	"                    "    "    : energía       "
45	"                    "    "    : máximo       "
46	"                    "    "    : entropía (coocurrencias)
47	"                    "    "    : energía       "
48	"                    "    "    : máximo       "
49	"                    "    "    : contraste     "
50	"                    "    "    : diferencia inversa (coocurrencias)

---



---

El primer experimento constaba de las 22 primeras variables y los máximos de las 6 primeras (la mayoría variables planimétricas más los momentos de orden 1,2,3,4 y sus desviaciones); el segundo experimento comprendía las variables texturales, incluidos los momentos y sus desviaciones (en total 30 variables); el tercer experimento englobaba a todas, en total cincuenta.

Cada grupo de características fué estudiado por imágenes divididas en dos clases: malignas y benignas (en total unas 2000 imágenes de núcleos y grupos y otras 2000 imágenes de sus fondos correspondientes). Cada imagen de los núcleos contenía un número variable de éstos y cada imagen de los grupos contenía un único grupo, ya fuese pequeño o grande. De las imágenes obtenidas de los campos seleccionados, se escogieron aquellas que reunían las mejores condiciones para el análisis (289 imágenes de grupos pequeños malignos, 150 de grupos pequeños benignos, 211 de núcleos benignos, 702 de núcleos malignos y 284 de grupos grandes benignos y malignos). Los resultados para las imágenes de los núcleos se estimaron como la media de los resultados de los núcleos que contenía.

También se realizó el estudio por casos (47 casos benignos, 95 malignos), cada caso constaba de varias imágenes de la lesión, benigna o maligna, proveniente de una sola paciente para obtener con ello: un mínimo de 25 núcleos,



---

entre 3 y 5 grupos pequeños y dos grupos grandes. Lo que hacía un total en el estudio de un mínimo de: 3550 núcleos y 850 grupos. Los resultados que se obtenían por caso se estimaron como la media de los obtenidos con las imágenes de dicho caso.

Los resultados de cada grupo de características fueron analizados usando una doble metodología: Análisis discriminante (BMDP 1990) y CART (CART 1984). La función de clasificación obtenida se testeó en ambos casos, en el caso del CART con una Validación Cruzada tomando K elementos para testear y el resto para construir el árbol (en nuestro caso K=10), repitiendo el test un determinado número de veces (n-K) y en el caso del Análisis Discriminante se utilizó un tercio de los casos para testear y los otros dos tercios para construir la función discriminante (Mosteller 1977). La proporción de casos acertados del grupo que valida nos da la medida de la bondad de los clasificadores que se obtuvo aplicando las fórmulas de la Eficacia (Et), Sensibilidad (S), Especificidad (Es), Valor predictivo (VPP) positivo y valor predictivo negativo (VPN) de los datos obtenidos en los test de clasificación (Chicharro y Burgos 1994).

$$\begin{aligned} Et &= 100 * (vp + vn) / total \\ S &= 100 * (vp / (vp + fn)) \\ Es &= 100 * (vn / (vn + fp)) \\ VPP &= 100 * (vp / (vp + fp)) \\ VPN &= 100 * (vn / (fn + vn)) \end{aligned}$$



---

Siendo:

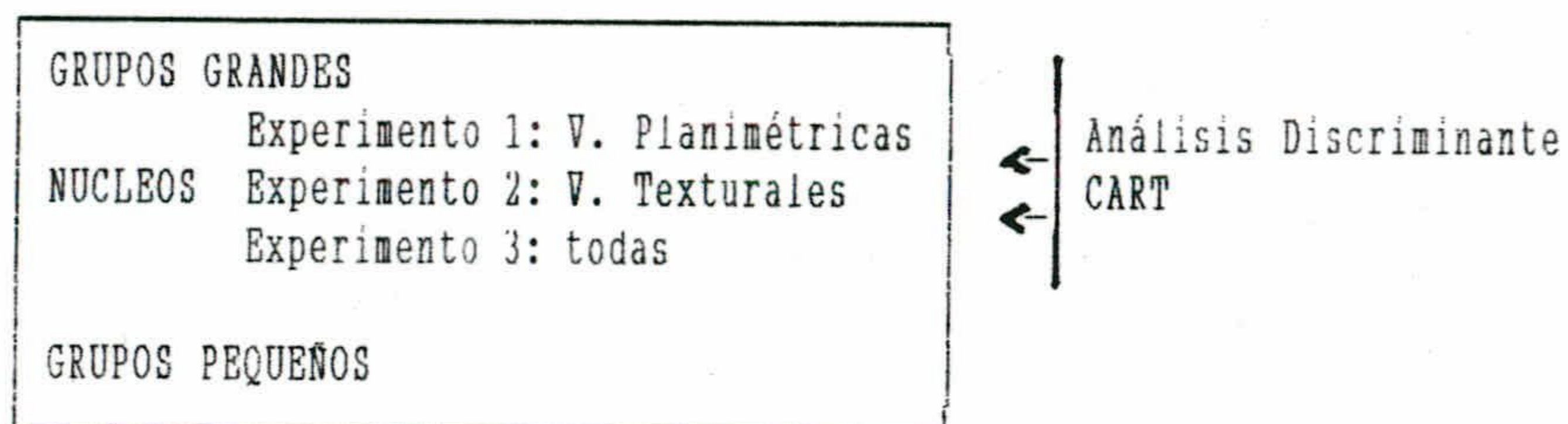
vp = verdaderos positivos (malignos en el test y verdaderos malignos).  
vn = verdaderos negativos (benignos " " benignos).  
fp = falsos positivos = 7 (malignos " " benignos).  
fn = falsos negativos = 5 (benignos " " malignos).

Estas fórmulas se aplicaron en los mejores clasificadores, aunque puede obtenerse en cualquier caso.

RESULTADOS POR IMÁGENES Y POR CASOS (benignos-malignos)

Variables

Clasificadores



Tras obtener los datos clínicos, todos los extendidos citológicos fueron examinados y señalados los campos adecuados para captar las imágenes.

La metodología del análisis de imagen incluye los siguientes procesos: EL CALIBRADO DE LA CÁMARA, LA SELECCIÓN Y CAPTACIÓN DE LAS IMÁGENES, LA SEGMENTACIÓN O UMBRALIZACIÓN DE LA ZONA DE INTERÉS Y EL ANÁLISIS.

## II-1 PROCESO DE CALIBRADO DE LA CAMARA

La operación de calibrado consiste en determinar, por comparación con una escala conocida (micrómetro objetivo) el valor (en micras por ejemplo) de cada división de la



---

escala medida en pixeles a través de una imagen captada del micrómetro. El algoritmo es muy simple, para cada objetivo (10X, 25X, 40X, 100X) hacer:

- Obtener una imagen del micrómetro (256X256)
- Repetir un número N de veces (N=30 por ejemplo)
- Calcular la longitud (en pixeles) de líneas trazadas sobre la imagen entre los puntos fijos del micrómetro a distancia D micras ( $\mu$ ). Sea  $L_i$  tal longitud.
- Calcular  $L = (1/N) \sum_{i=1}^N L_i$

La escala resultante es  $1\mu = \text{Round}(1/D)$  pixeles.

Este proceso de calibración nos permitirá, en todos los algoritmos que conlleven medidas, poder dar las salidas de éstas en unidades interpretables desde el punto de vista biomédico. El software Visilog cuenta con funciones para hacer el calibrado automáticamente. En nuestro caso todas las medidas se han dado en pixels pues nuestro interés se centraba en la asignación de clases con independencia de la medida y los pixels constituyen la medida directamente utilizada por el analizador y por el software desarrollado.

## **II-2 PROCESO DE CAPTACION DE LAS IMAGENES**

### **II-2-1 Camara de video**

Para la captación de imágenes se ha usado una cámara de video tipo CCD adaptada al microscopio. Si notamos  $s(x,y)$  y  $E(x,y)$  las señales de salida y entrada del video respectivamente tenemos (Levine 1980):



---

$$S(x,y) = K * E(x,y)^\delta$$

Con  $K=$ Constante y  $\delta$  un parámetro dependiente del tipo de cámara. En nuestro caso  $\delta \approx 1$  con lo que puede considerarse que la respuesta electrónica  $s(x,y)$  es esencialmente una función lineal de la entrada o equivalente. Puede suponerse la existencia de una relación lineal entre la luminosidad medida y la verdadera en cada pixel:

$$\text{LUMINOSIDAD OBSERVADA} = \text{LUMINOSIDAD VERDADERA} * \text{CTE.}$$

#### II-2-2 Depuración del proceso de captación

Una de las primeras aplicaciones implementadas en el software es la depuración del proceso de captación. Existen numerosas fuentes de distorsión de la imagen provenientes unas de la falta de resolución de la cámara de video y tarjeta digitalizadora, otras de la falta de uniformidad en la distribución de la luz y posibles puntos de suciedad. Para eliminar éstas distorsiones se toma una imagen del fondo correspondiente a cada toma de imagen, es decir sustituyendo ésta por un cristal limpio (con lo cual tendría las mismas condiciones de captación). Utilizando ambas imágenes la función de depuración obtiene una nueva aplicando la ecuación:

$$R_j = 255 * \text{minimo} \{1, B*(I_j/F_j) \} \quad (1)$$



---

Siendo:

- R, la imagen transformada
- I, la imagen inicial
- F, la imagen fondo correspondiente a I.
- B, constante,  $B < 1$  (un valor aconsejable para B se encontró que era  $B=0,75$ ).

Las imagenes que se consiguen aplicando (1) tienen un fondo casi constante con niveles de grises altos lo que permitirá una mejor selección de la zona de interés. La razón de multiplicar por B es para trasladar el histograma. Si no multiplicamos por una constante menor que uno, perdemos la información de los bordes de la imagen y de otras estructuras presentes ya que, por efecto de la retracción del flujo de luz al cruzarlas, el cociente  $I_j/F_j$  es mayor que 1, cuando debería ser menor ya que los valores de  $I_j$ , que representan niveles de grises son menores que los de  $F_j$ . La selección del valor de B se obtuvo por experimentación de prueba y error. Con ésta transformación los parámetros estadísticos más usuales en las imágenes: moda, media y rango, se nos modifican de la siguiente manera:

$$\text{Moda} \approx 255 * \min\{1, B\} = 255 * B = 255 * 0,75 = 191$$

La moda en las imágenes transformadas tiene un valor que fluctúa alrededor de 191, independientemente de la moda de las imágenes iniciales.

$$\text{Media} = 255 * \min\{1, B*(I/F)\} = \begin{cases} 255 & \text{si } B*(I/F) > 1 \\ 255*B*(I/F) & \text{si } B*(I/F) < 1 \end{cases}$$



---

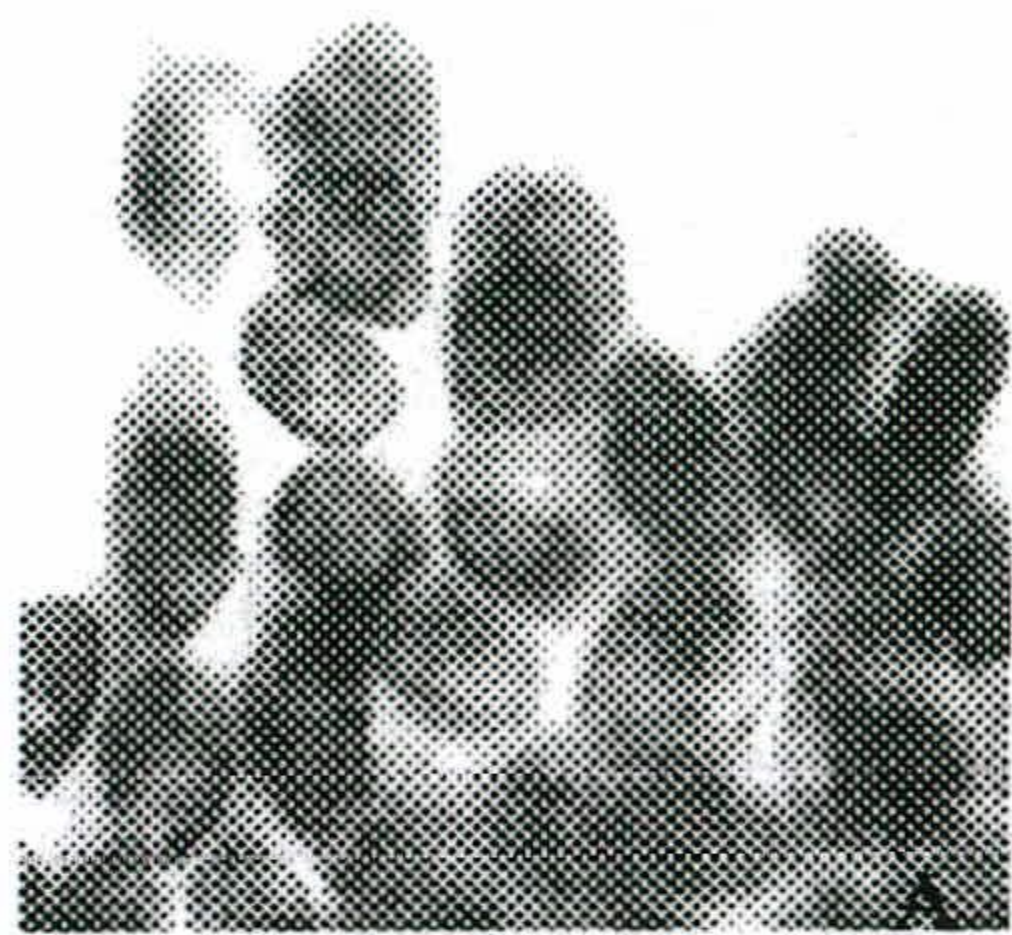
Siendo  $I/F$  la media de los niveles en la imagen inicial y la media de los niveles en la imagen del fondo y, puesto que los distintos valores de dicho cociente son muy parecidos, se obtienen imágenes transformadas con media similares, independientemente de las condiciones de captación.

Finalmente, la aplicación de (1) lleva en la práctica a la ampliación del rango de niveles de grises en la imagen transformada con respecto a la inicial, ya que en teoría el rango de las imágenes transformadas puede llegar a ser  $(0, 255)$  puesto que si en algún pixel  $\beta \cdot (I/F) > 1$  se tendrá  $R_j = 255$  y en el caso de que  $\beta \cdot (I/F) < 1$ , entonces  $R_j = (255 \cdot \beta) \cdot (I_j/F_j)$  de forma que si en algún pixel  $I_j=0$  con  $F_j$  distinto de 0 se tendrá  $R_j=0$ .

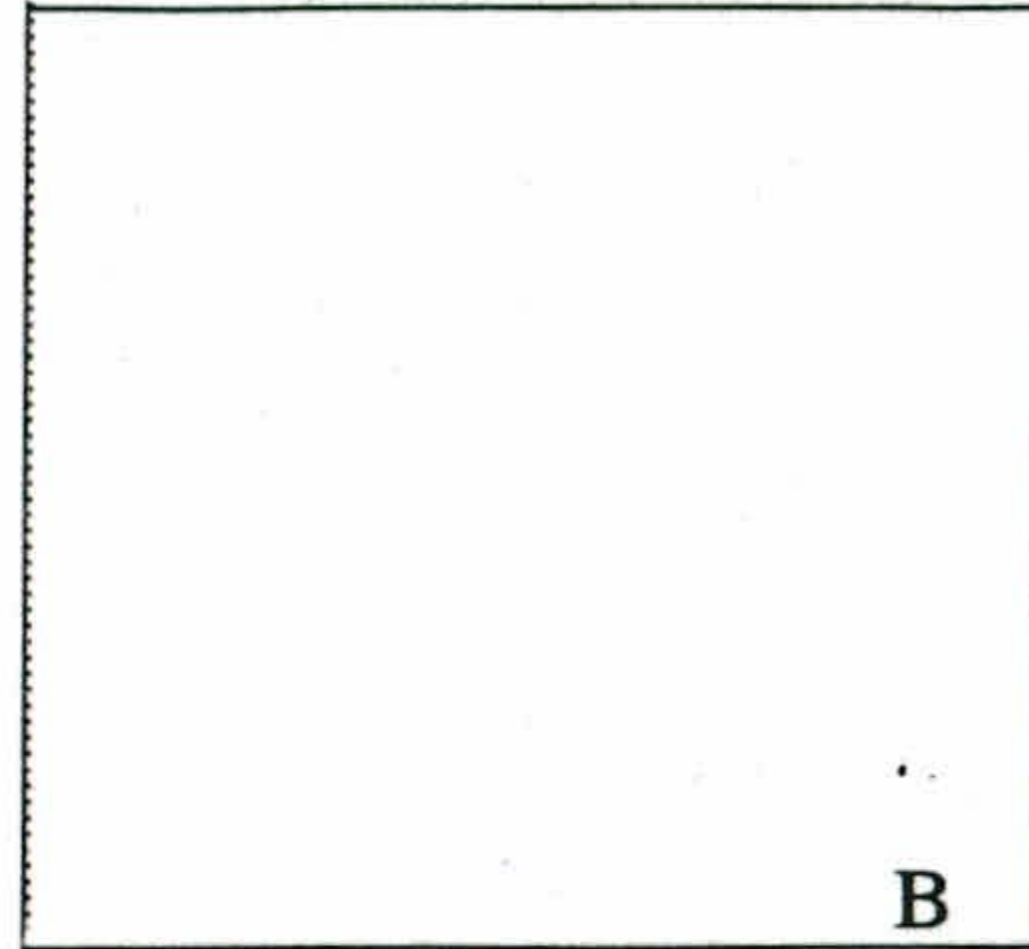
En la Figura 7 puede verse ejemplos de imágenes iniciales, sus fondos respectivos y las imágenes transformadas.



Figura 7

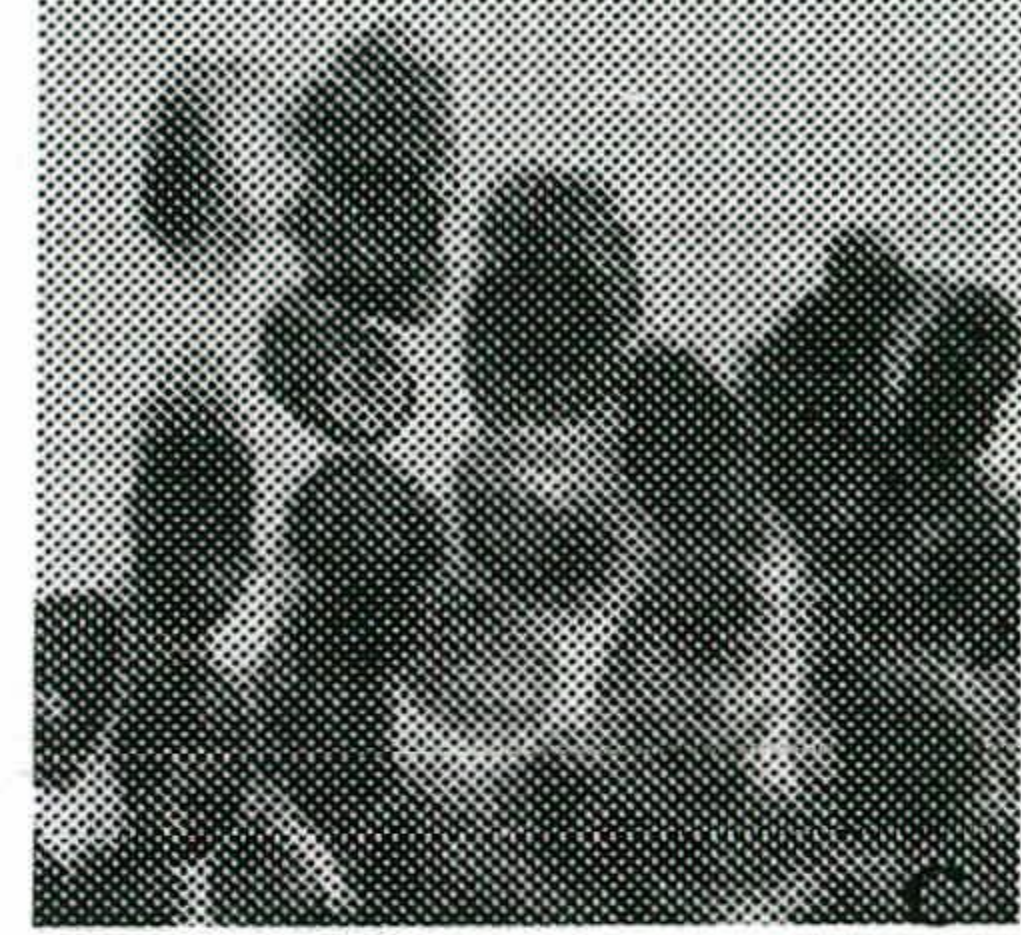


clm00340

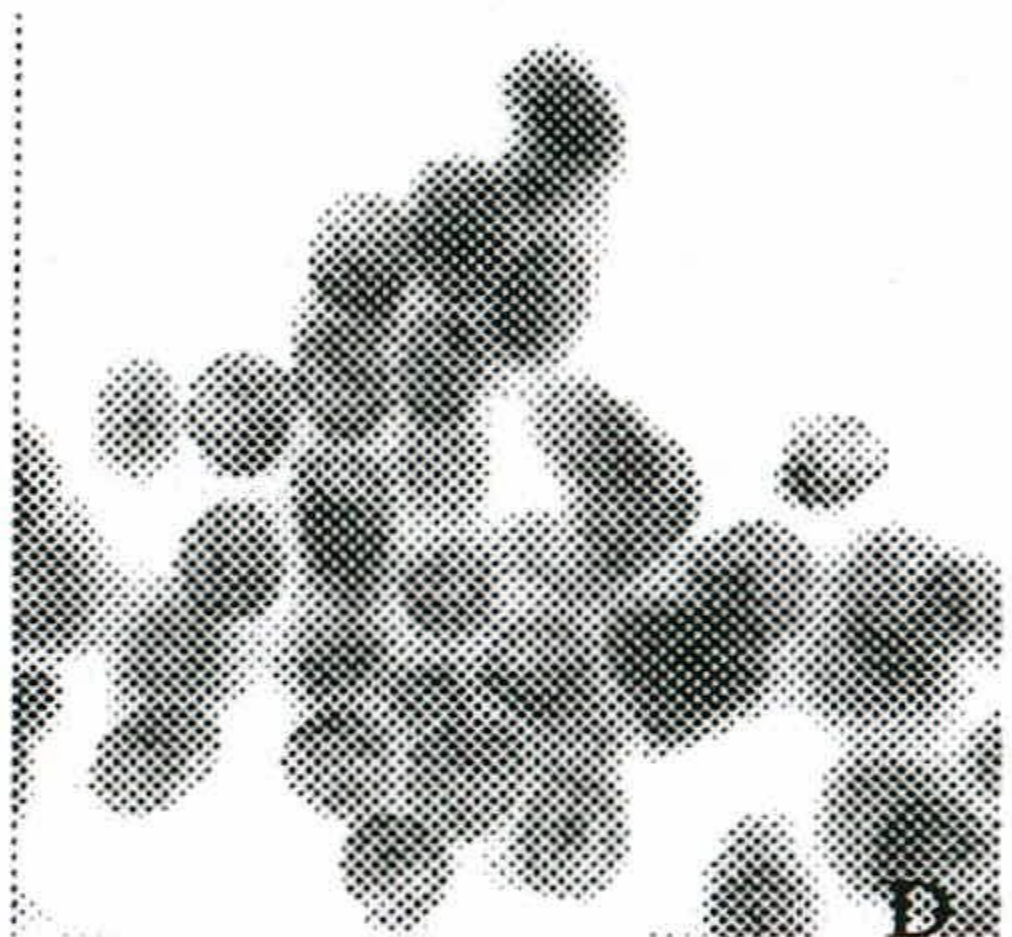


B

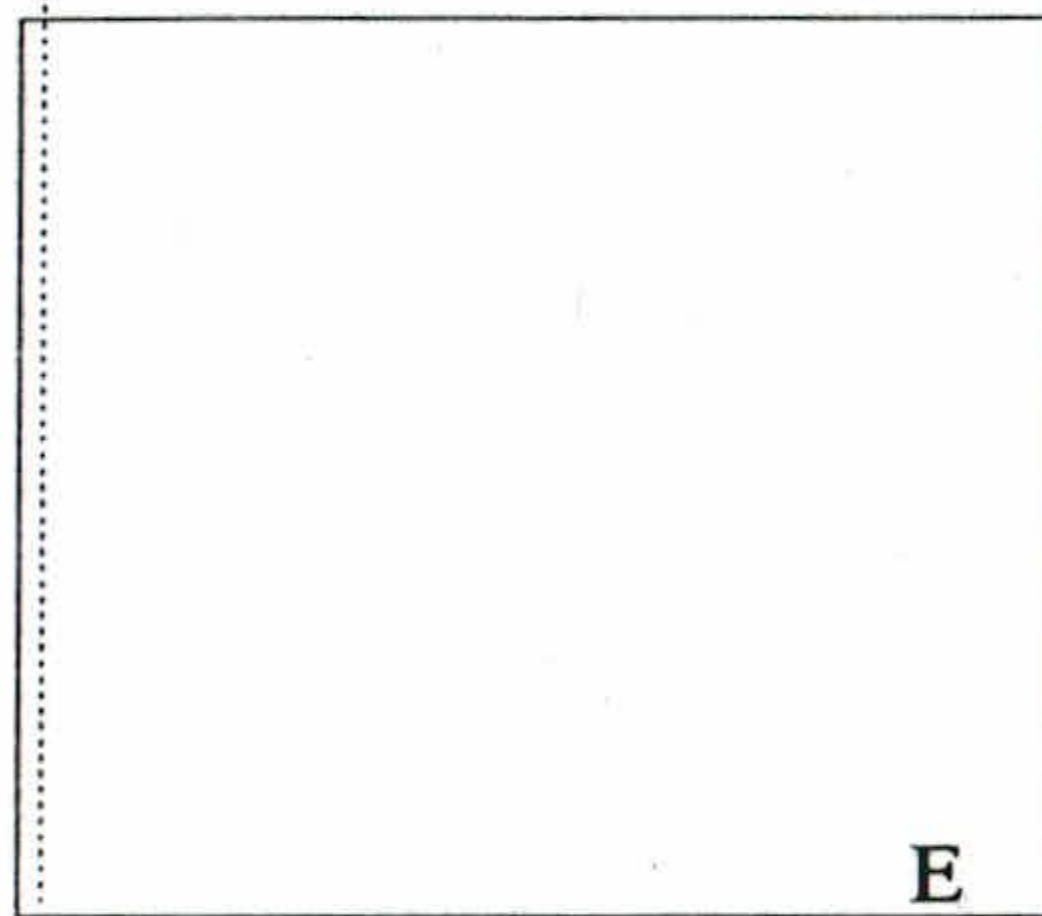
cfm00340



crm00340

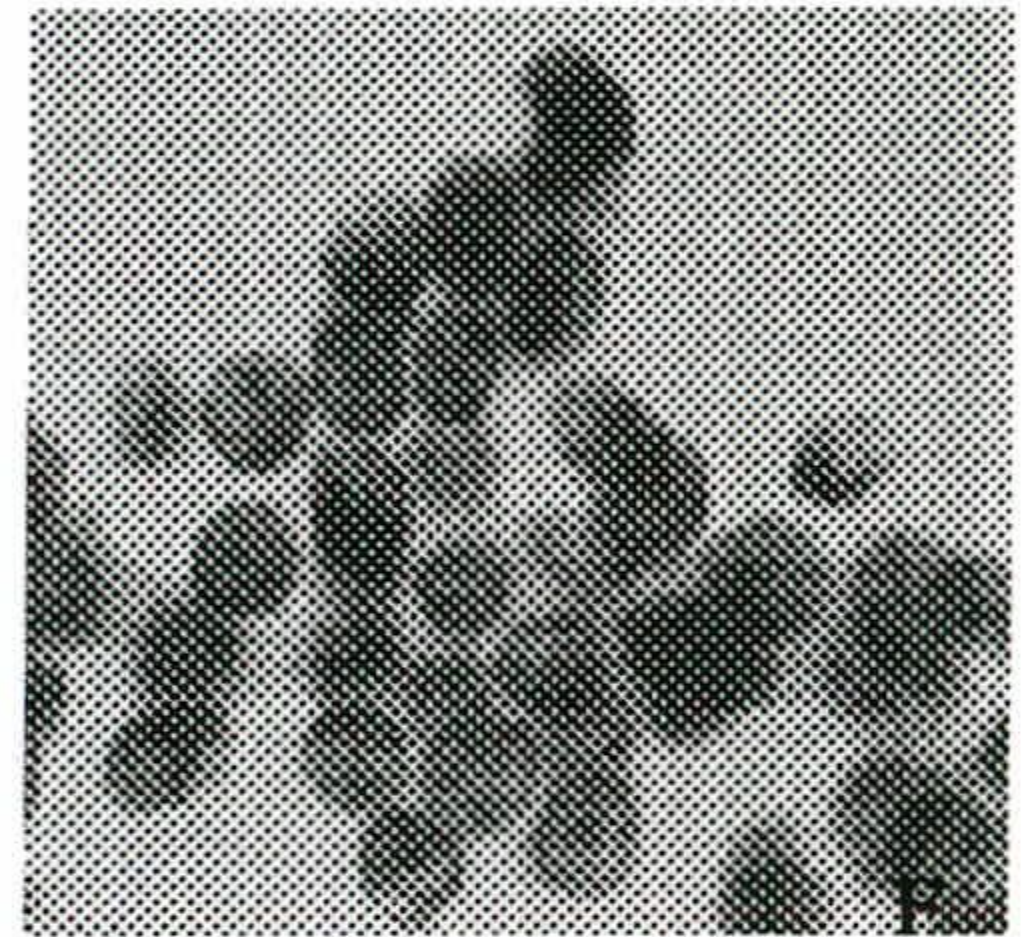


clb01620

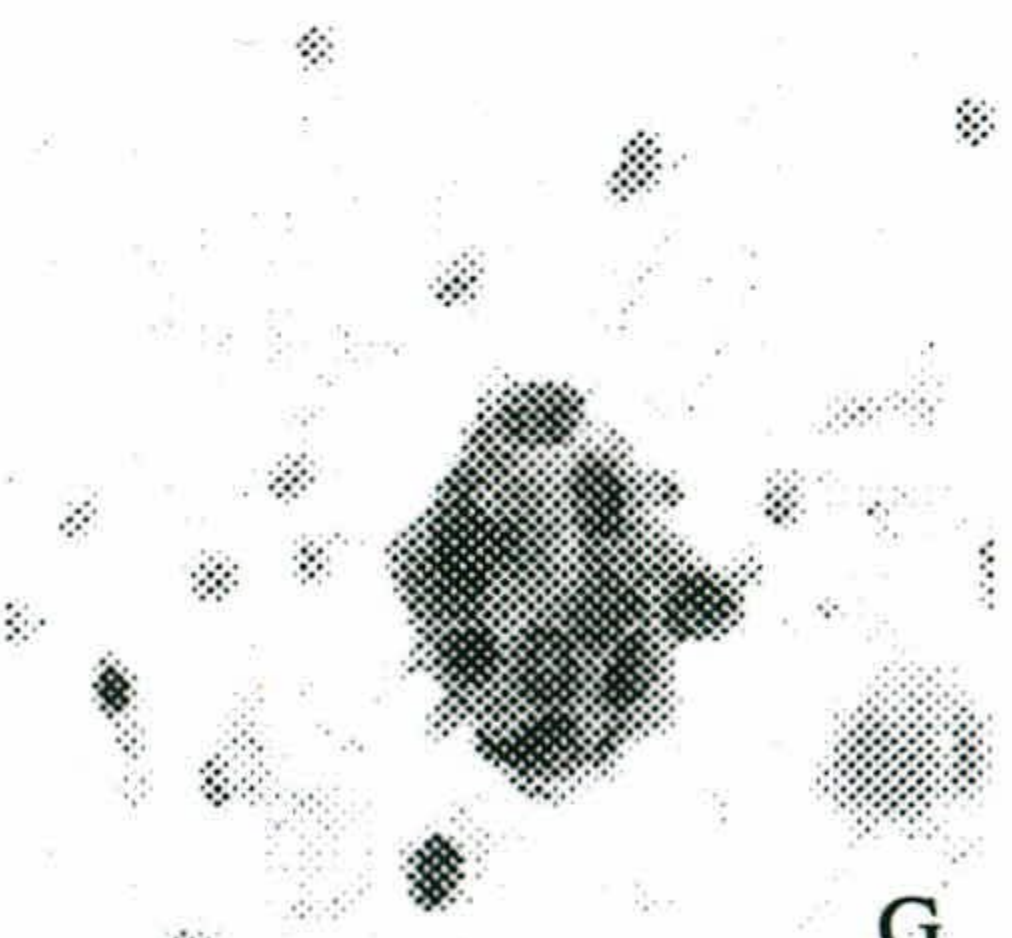


E

cfb01620

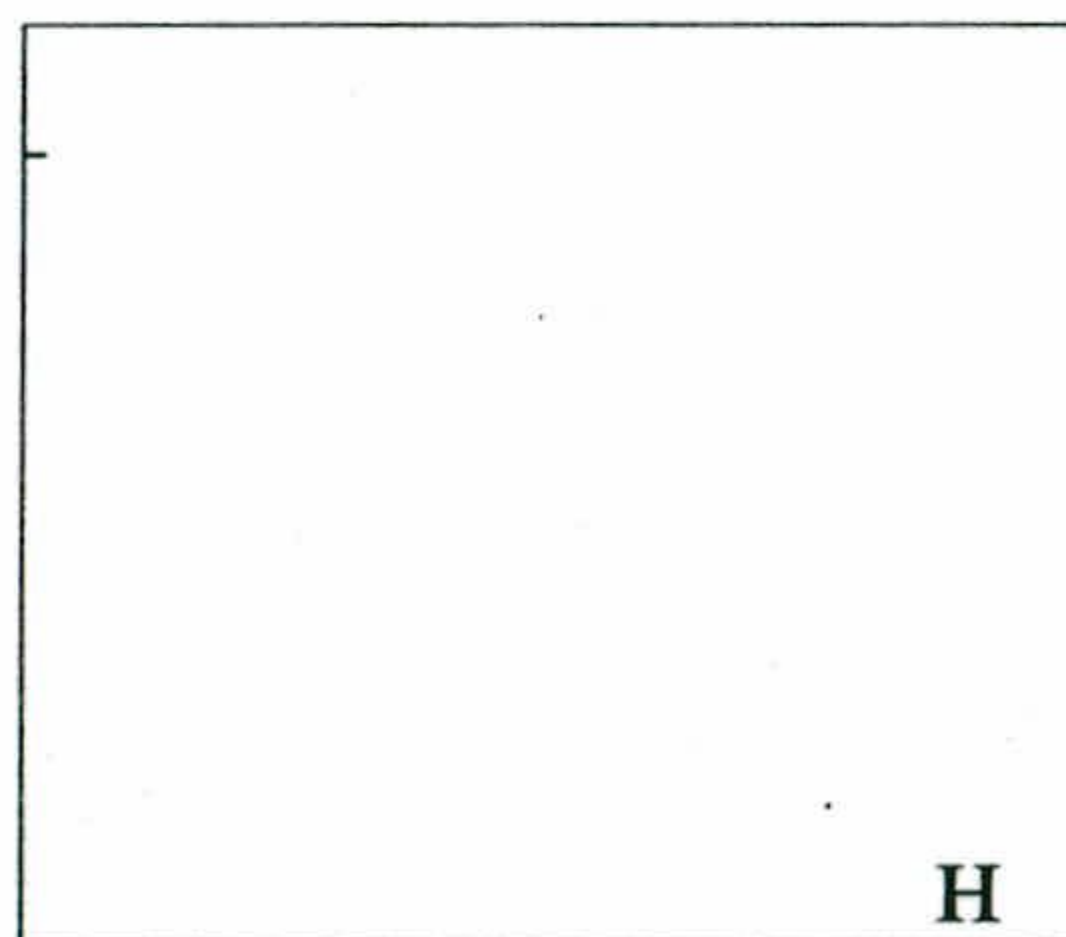


crb01620



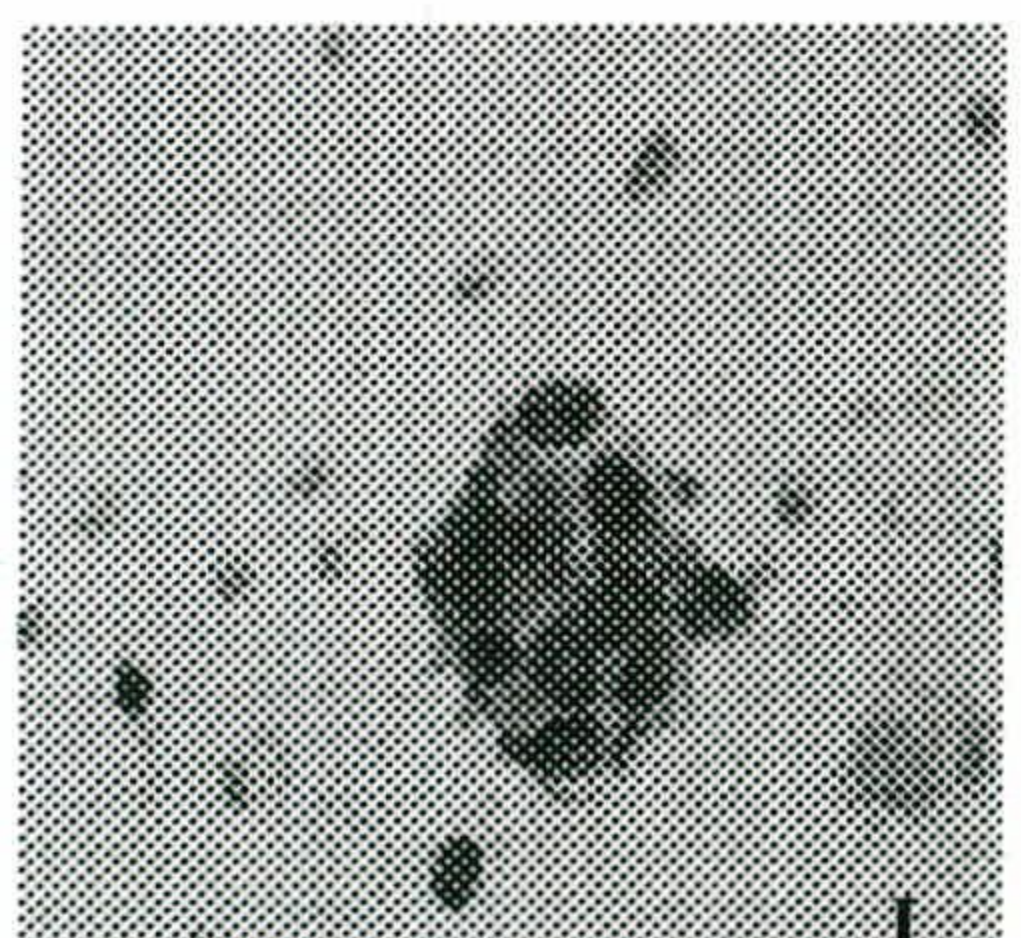
plm00334

G



H

pfm00334



prm00334

I



---

## II-3 PROCESO DE SELECCION DE LAS IMAGENES

Las imágenes de los grupos grandes se obtuvieron con un objetivo de 10x o 5x y apertura numérica (AN) de 0,25 y 0,12 respectivamente. Las imágenes de los grupos pequeños con un objetivo de 40x con AN de 0,65 y la de los núcleos con 100x y 1,25 de AN. Se tomó en las mismas condiciones una imagen del fondo con un porta limpio por cada imagen inicial con el objeto de obtener una imagen final libre de la suciedad y defectos de iluminación inherentes al sistema.

Para analizar los grupos grandes se escogieron los dos grupos más grandes que existían entre todos los portas obtenidos en cada caso.

En el análisis de los grupos pequeños se esperaba un sesgo por defecto en la obtención de la relación entre el área del grupo y de los núcleos debido al solapamiento de los núcleos (cabe esperar, por ello, que sea mayor en las lesiones malignas). Para minimizar éste error se tomaron imágenes de 3 a 5 grupos que, en un rastreo previo, se eligieron por su menor solapamiento, mayor nitidez de contornos nucleares, citoplasmáticos y cromatínicos, escasa degeneración y desecación. Para las variables de los núcleos se obtuvieron imágenes que contuviesen en total un mínimo de 25 núcleos por caso, siempre núcleos integrantes de los grupos. En la selección se siguieron los mismos criterios



---

anteriormente expuestos.

## **II-4 SEGMENTACION DE LA ZONA DE INTERES**

Para el estudio de las variables de los grupos pequeños se utilizaron dos programas: una MACRO (conjunto de funciones encadenadas entre sí que realizan una tarea) de tratamiento y análisis semiautomático realizada usando el programa VISILOG (1988), para la obtención de la relación entre AN/AC de los grupos y para el desorden de los núcleos integrantes, fué utilizado el software de análisis automático basado en distancias desarrollado al efecto. Para el estudio de las variables nucleares: área, perímetro, ejes, factores de forma...etc. y análisis de los grupos grandes fué utilizado también software específico.

### **II-4-1 MACRO PARA LA RELACION NUCLEO-CITOPLASMATICA EN LOS GRUPOS**

La relación AN/AC de las células es un rasgo citológico importante para el diagnóstico ya que se altera en la malignidad. Medir ésta relación en las células individualizadas de los extendidos malignos y benignos no era posible ya que, si bien, en los extendidos malignos las células suelen desgajarse y encontrarse muchas de ellas sueltas, no ocurre ésto en los extendidos benignos donde la cohesión en los grupos es la norma lo que obligaría a



---

utilizar segmentaciones manuales, procedimiento contrario al objetivo de éste trabajo. Por ello se intentó una aproximación realizando la relación de áreas sobre el grupo en su conjunto: área total del grupo dividido por el área total de los núcleos. Estos grupos debían tener un mínimo de 3 núcleos, un máximo de 25 núcleos y un tamaño que pudiese ser captado con holgura por la cámara de vídeo con un objetivo de 40X.

La macro utilizada se basa en técnicas de Morfología Matemática y consta fundamentalmente de: una función de NORMALIZACIÓN que realiza la expansión de los niveles de grises en los que se distribuye la imagen, a toda la gama de que se dispone (256 niveles de grises). Con ello se consigue un mayor contraste entre estructuras lo que permite una mejor BINARIZACIÓN O UMBRALIZACIÓN interactiva. Esta función reduce a dos los niveles de grises, el 1 contendrá la zona de interés para el análisis y el 0 todas las restantes. El proceso es el siguiente:

- Primero se binariza la imagen eligiendo interactivamente el umbral que mejor segmente el objeto, en éste caso el grupo total en cuestión.
- Segmentación de forma manual, interactuando con la imagen y utilizando la función de pintar para eliminar aquellas zonas que consideremos ajenas al grupo y que hayan quedado incluidas en él por similitud de niveles de grises. Estos artefactos suelen producirlos hematíes o células del estroma adheridos al grupo y zonas de sombra.
- A continuación se aplica la función de etiquetar. Esta función va a poner diferentes etiquetas representadas en pseudocolor a todos aquellos objetos que formen una componente conexa diferenciada. Seguidamente se elige el grupo a analizar.



---

■ La segunda parte es la binarización de los núcleos y la aplicación de una función de XOR o resta entre el área del grupo total (G) y el área de los núcleos (N) integrantes, obteniendo el área del grupo libre de los núcleos (C) o área del citoplasma.

■ Finalmente se obtiene el área de los núcleos (N) por la resta al área del grupo, el área del citoplasma. Con ello obtenemos dos índices de relación:

$$I = N/G \quad \text{y} \quad I' = N/C$$

En la Figura 8 podemos ver imágenes secuenciadas en la aplicación de la macro.



---

Figura 8a y 8b: imagen digitalizada de grupo pequeño (400 aumentos) y su histograma de niveles de gris

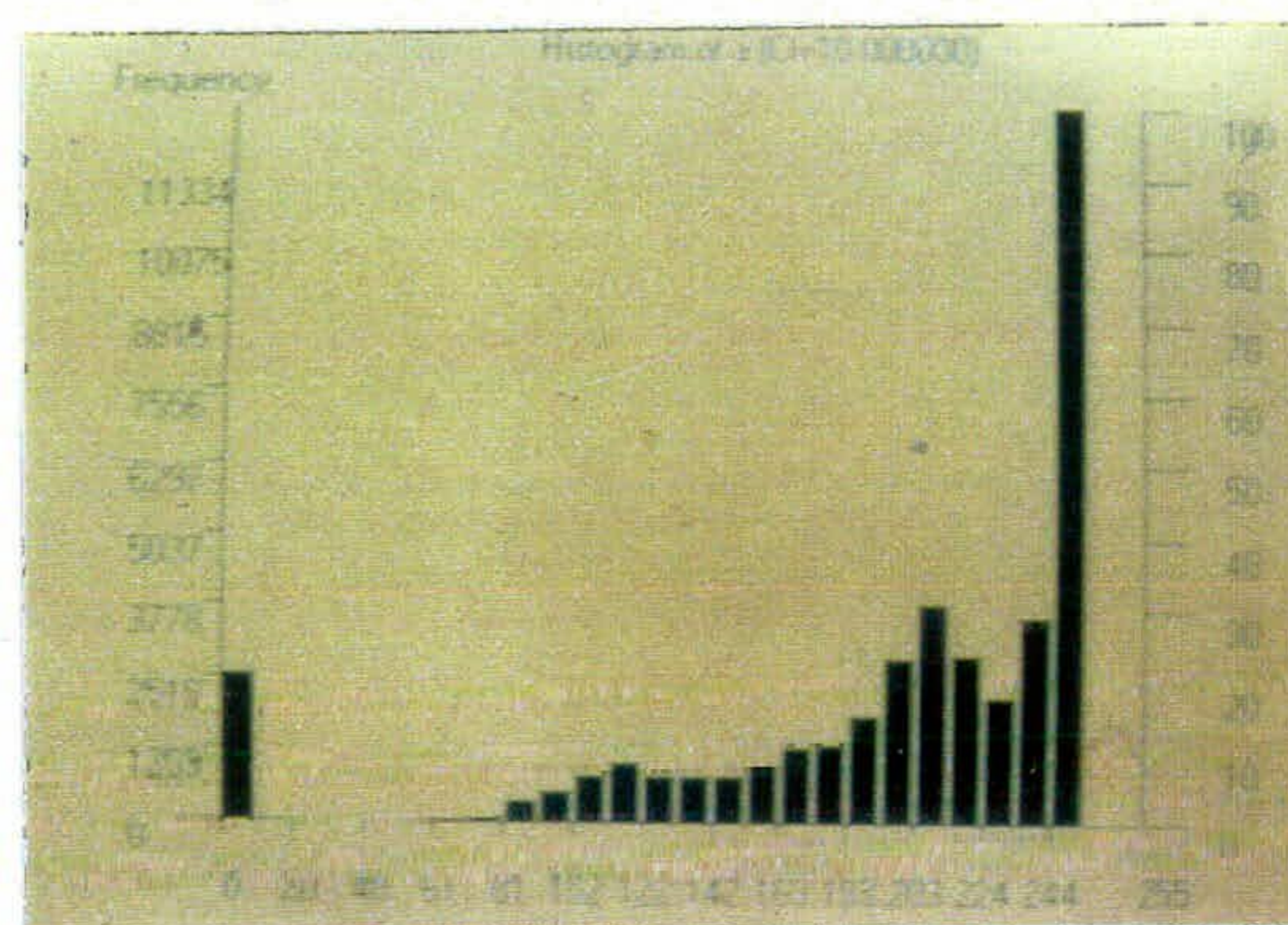


Figura 8c: Binarizado del grupo completo

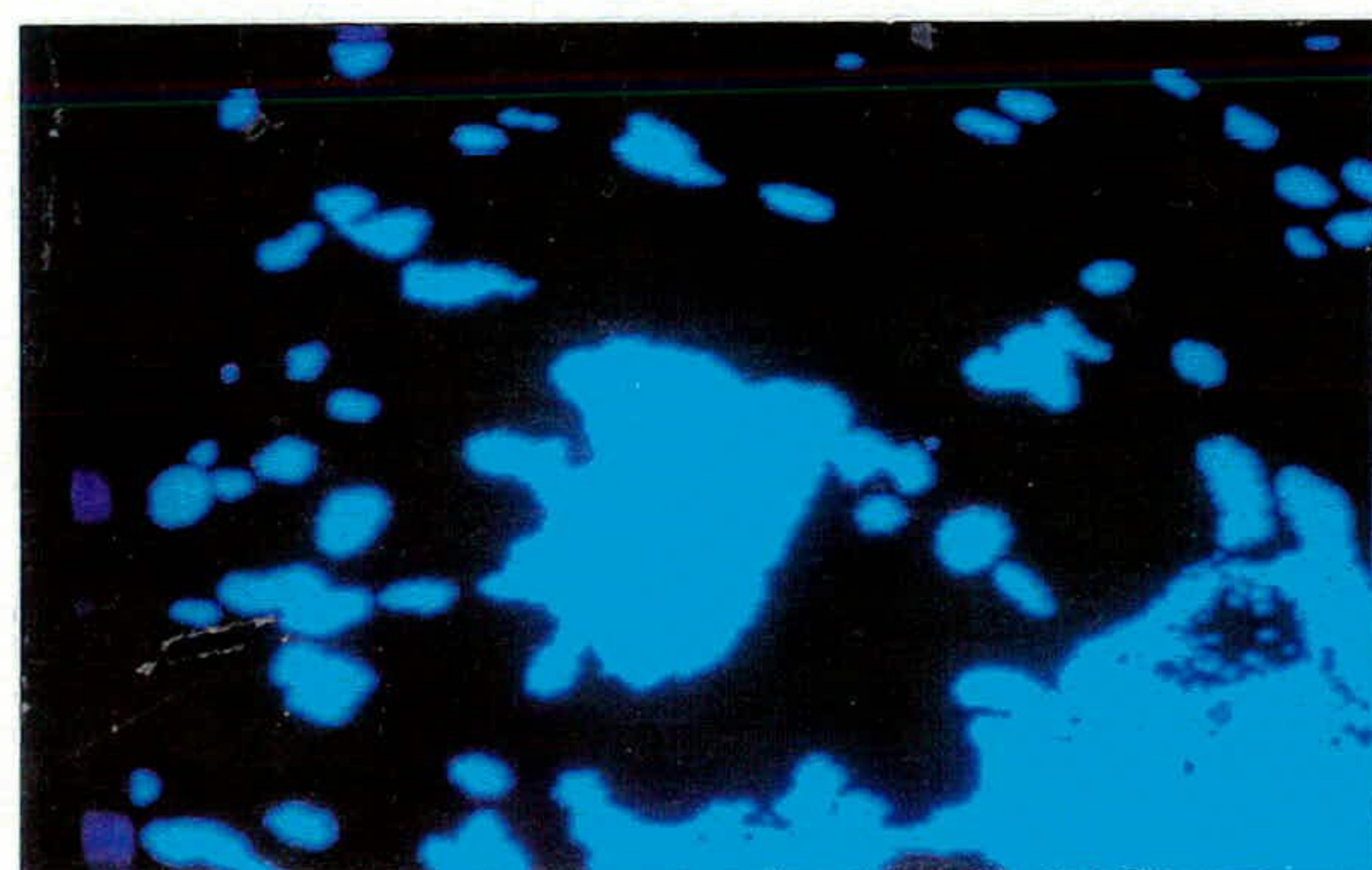


Figura 8d: Binarizado de los núcleos

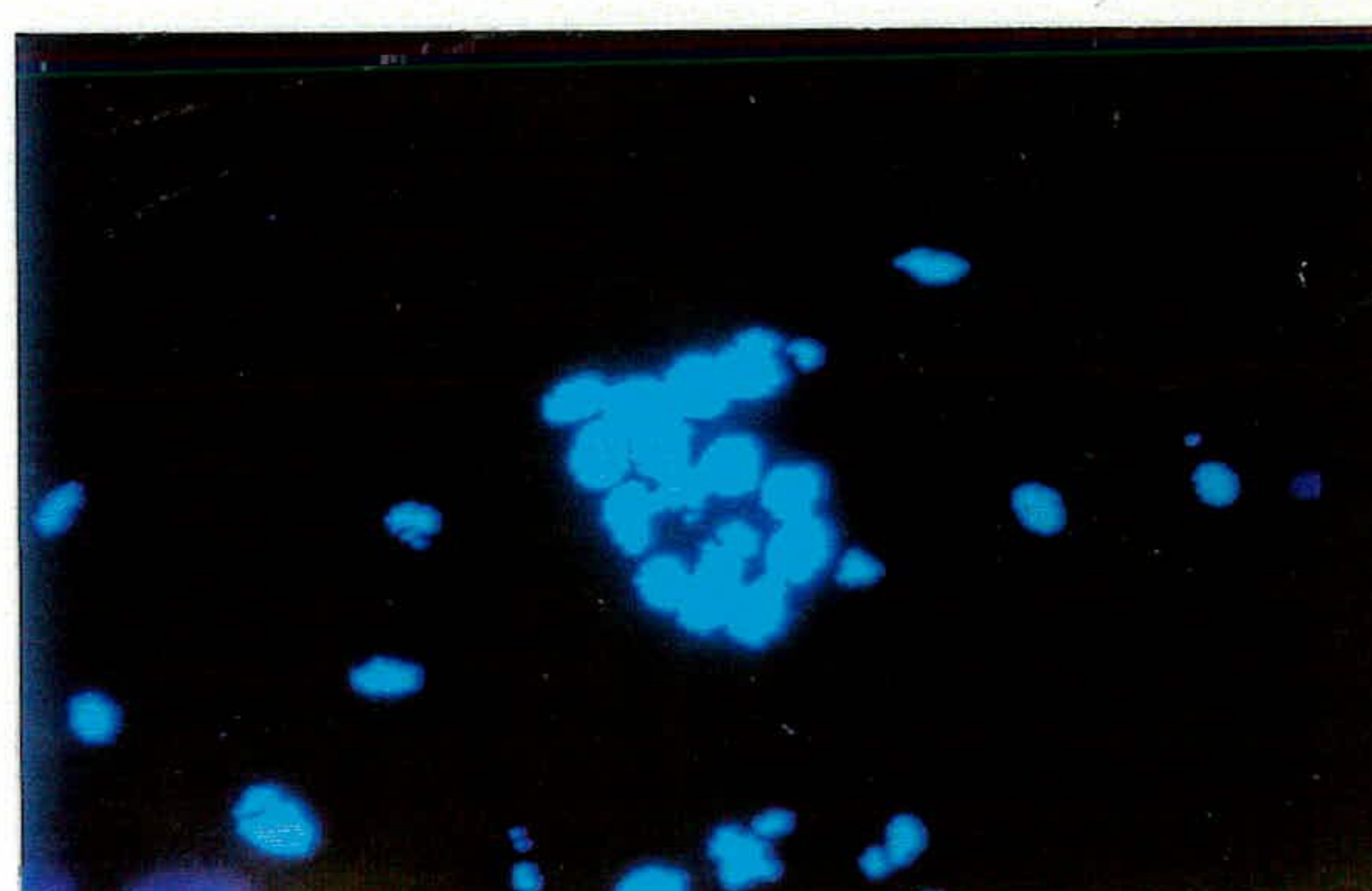


Figura 8e: Función de xor o resta de la figura 8c y 8d: área citoplasmática del grupo





---

## II-4-2 DESCRIPCION DE LAS FUNCIONES DEL SOFTWARE ESPECIFICO DE ANALISIS AUTOMATICO DE LOS GRUPOS PEQUEÑOS BASADO EN DISTANCIAS NUCLEARES: Creación de grafos planares.

La idea es analizar una propiedad espacial que puede resultar discriminatoria para los distintos tipos citológicos: la distancia internuclear, que aparentemente resulta más irregular en las células malignas que en las benignas. Dicho de otra forma, se intentará hacer objetivo el criterio del experto de usar el desorden espacial en la distribución de las células en los grupos, como una medida de diferenciación entre malignidad y benignidad.

### Construcción de grafos planares

La idea que se usó para resolver el problema tué la de utilizar un grafo planar que conecte el centro de cada nucleo con los centros de los nucleos más próximos del mismo grupo de células. Se analizó el grafo resultante a través de 4 variables: LA MEDIA Y DESVIACIÓN DE LAS DISTANCIAS Y EL MÁXIMO Y MÍNIMO DE LAS MISMAS. El grafo planar tiene, por tanto, como nodos los centros de las células  $C_i$  y como peso de las aristas las distancias internodos  $d_{i,j} = d(C_i, C_j)$ .

Existen muchas técnicas para la construcción de grafos planares que deban de cumplir ciertas propiedades. En nuestro caso el grafo debería de ser planar con sus nodos



---

conectados únicamente a sus vecinos más próximos. El más conocido de los algoritmos está basado en la construcción previa del diagrama de Voronoi. Aquí se diseñó un algoritmo alternativo al clásico (Preparata 1987) que, aprovechándose de propiedades geométricas existentes entre los nodos del grafo y los polígonos de Voronoi, permitía obtener un grafo construido a partir de triángulos que cumplen la propiedad de ser lo más regulares posibles.

Una vez señalados interactivamente los centros nucleares en el grupo de células epiteliales, que constituiría los vértices o nodos de los triángulos que compondrá el grafo, se aplica el algoritmo automático desarrollado (Algoritmo Geométrico de Construcción). Este consta de los siguientes pasos:

1 Para cada nodo construye todas las rectas normales a los segmentos de recta que unen ese nodo con el resto.

2 Selecciona las intersecciones que forman parte del diagrama de Voronoi como sigue: Construye un segmento de recta desde el nodo (al que vamos a calcular su polígono asociado) a una intersección cualquiera. Se producen entonces dos situaciones:

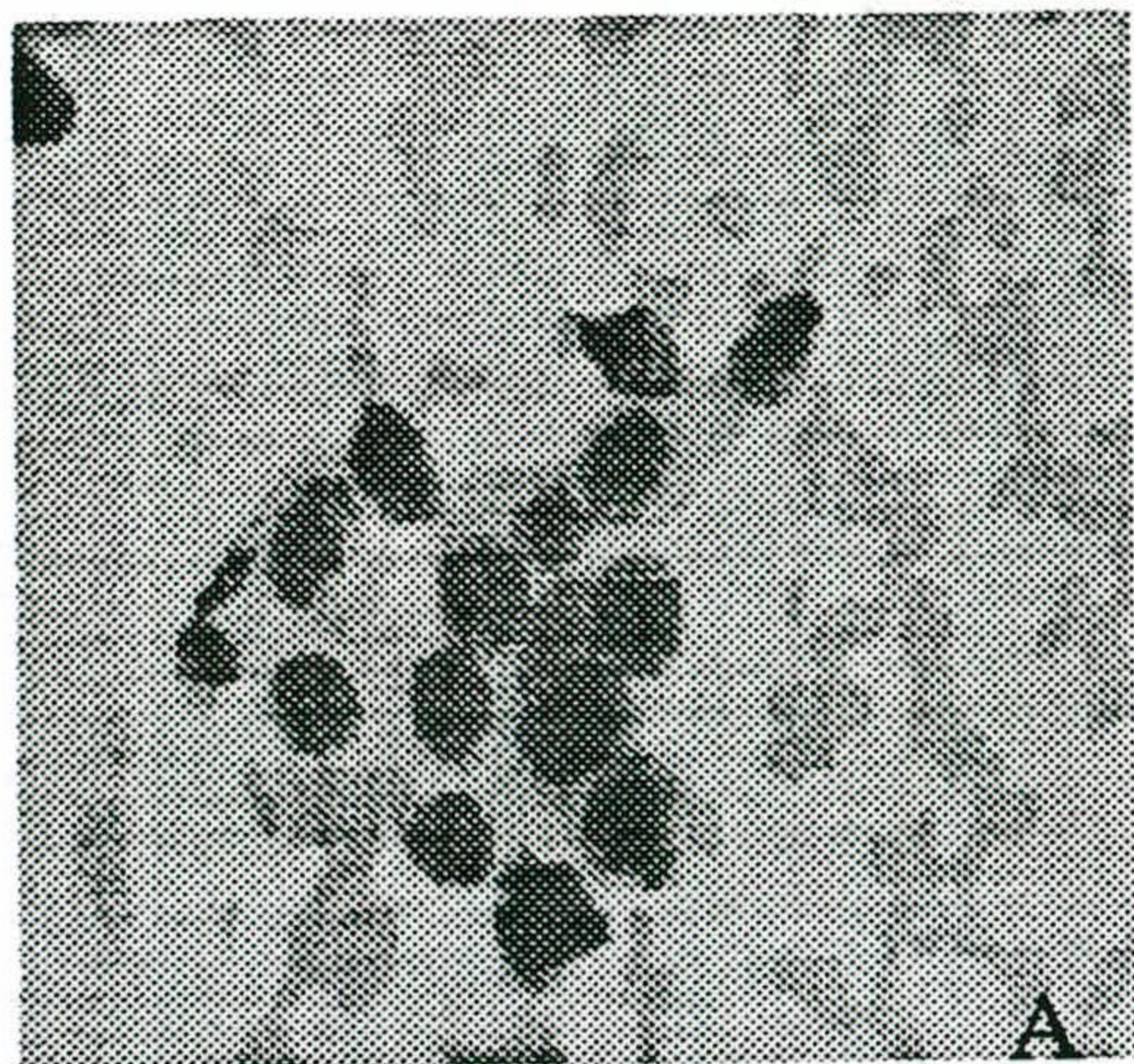
A-Si el segmento no interseca con ninguna recta existente, entonces el punto extremo del segmento es un vértice del polígono de Voronoi asociado al nodo.

B-Si el segmento interseca con alguna recta existente entonces ignora el punto extremo del segmento.

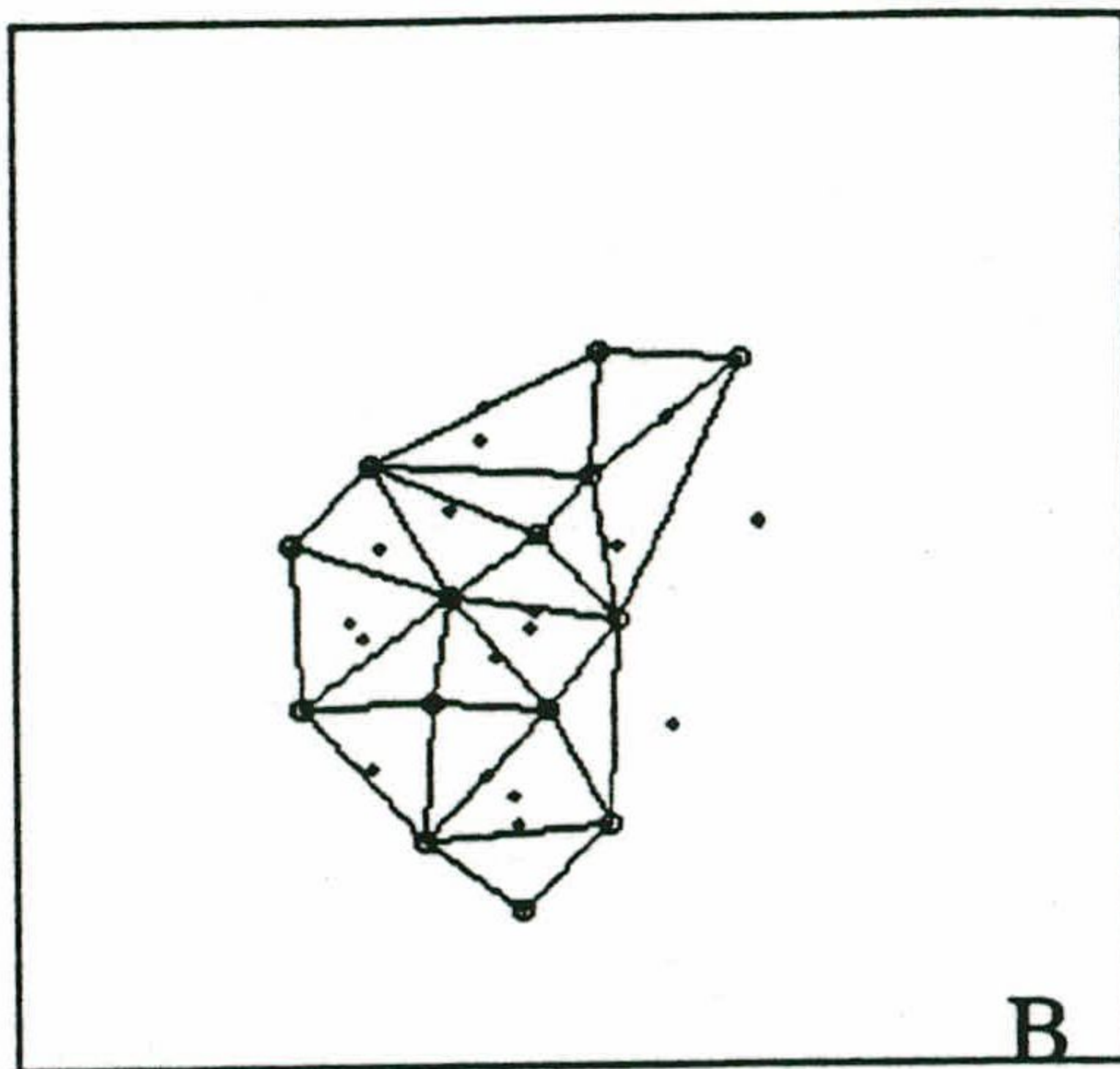
3 Triangula el grafo construido tomando para cada vértice considerado su unión a los tres y sólo tres nodos del grafo a los que está asociado.

Se pueden ver algunos ejemplos de grafos en la Figura 9.

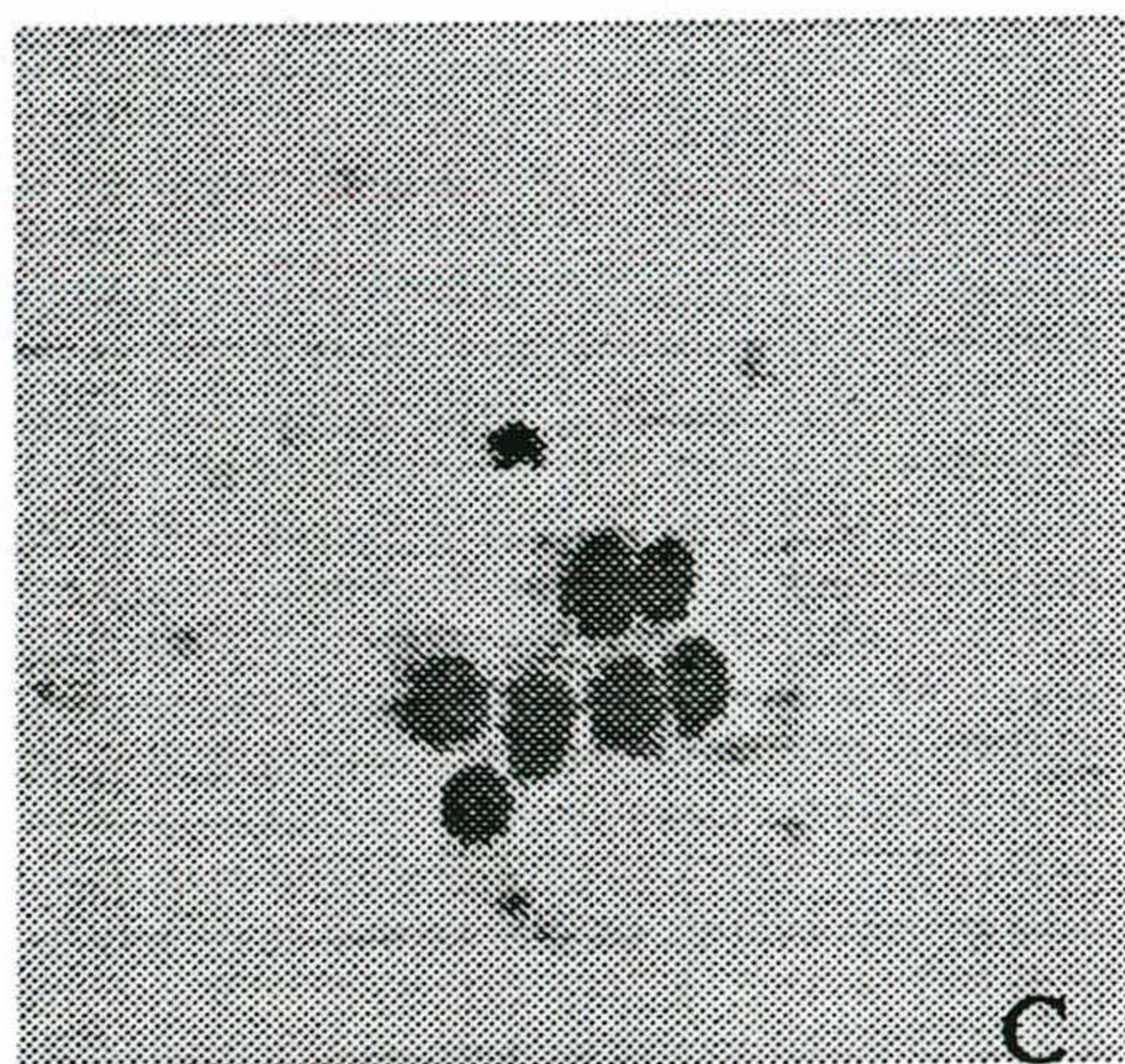




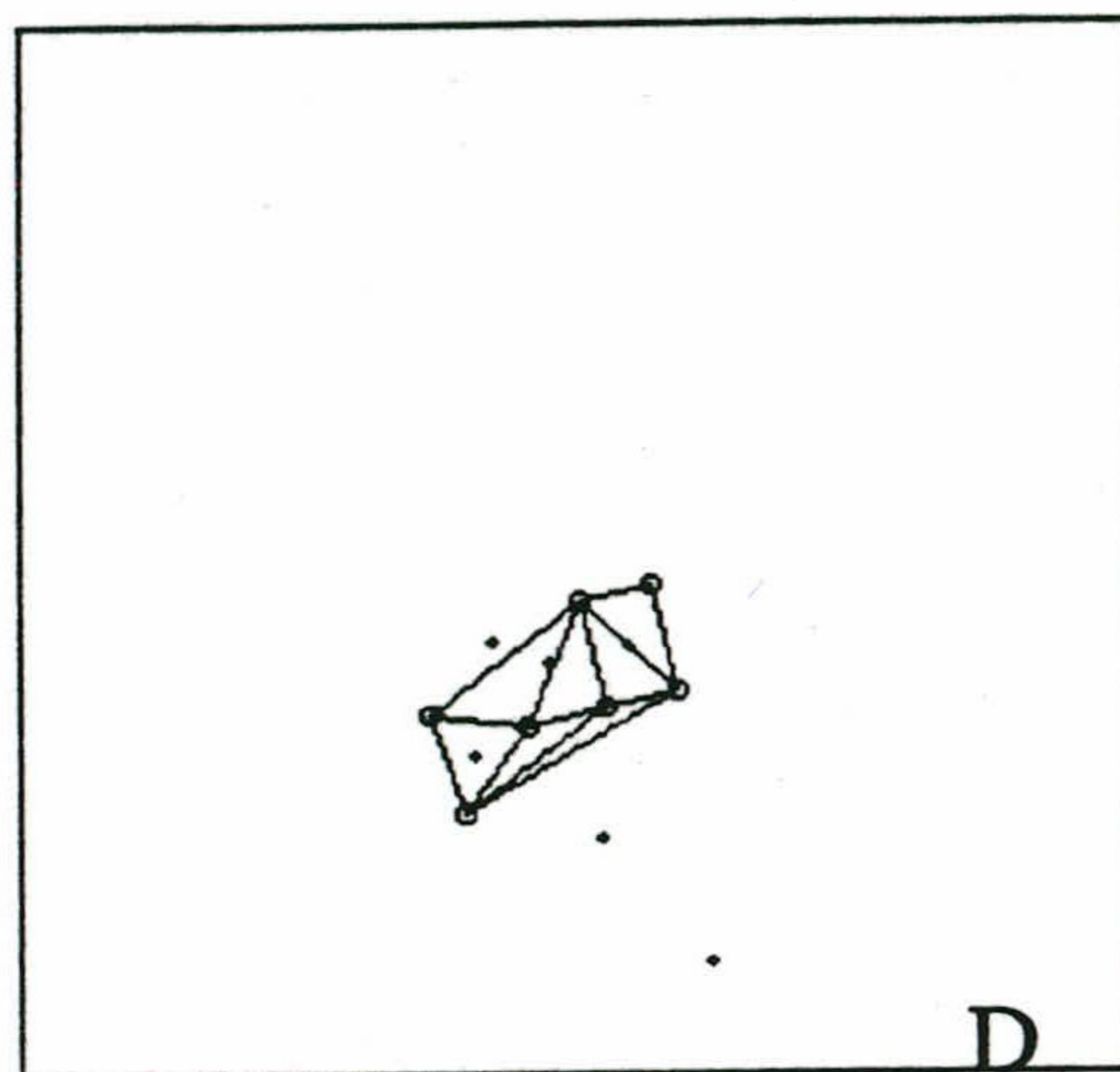
prm05064



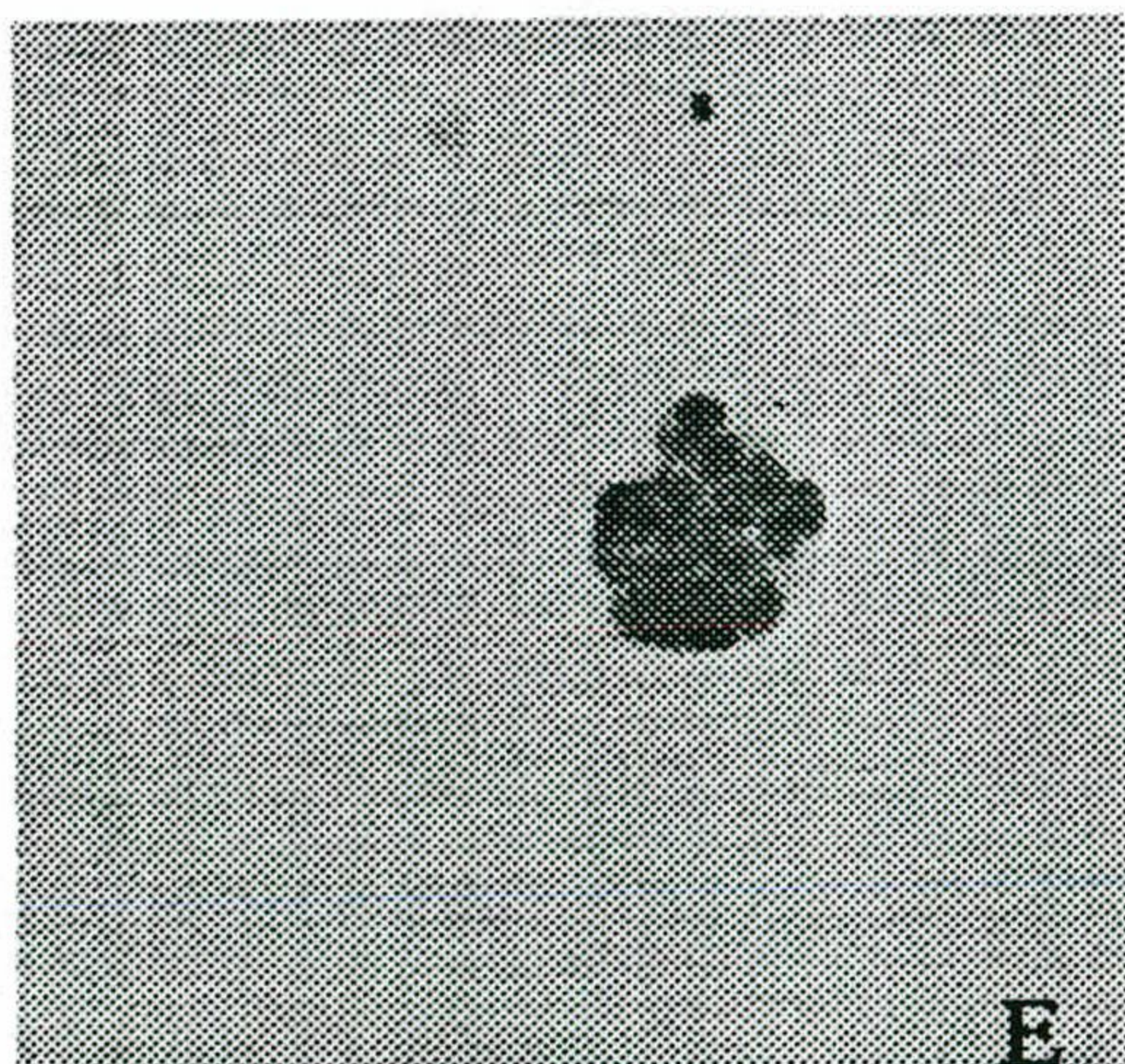
B



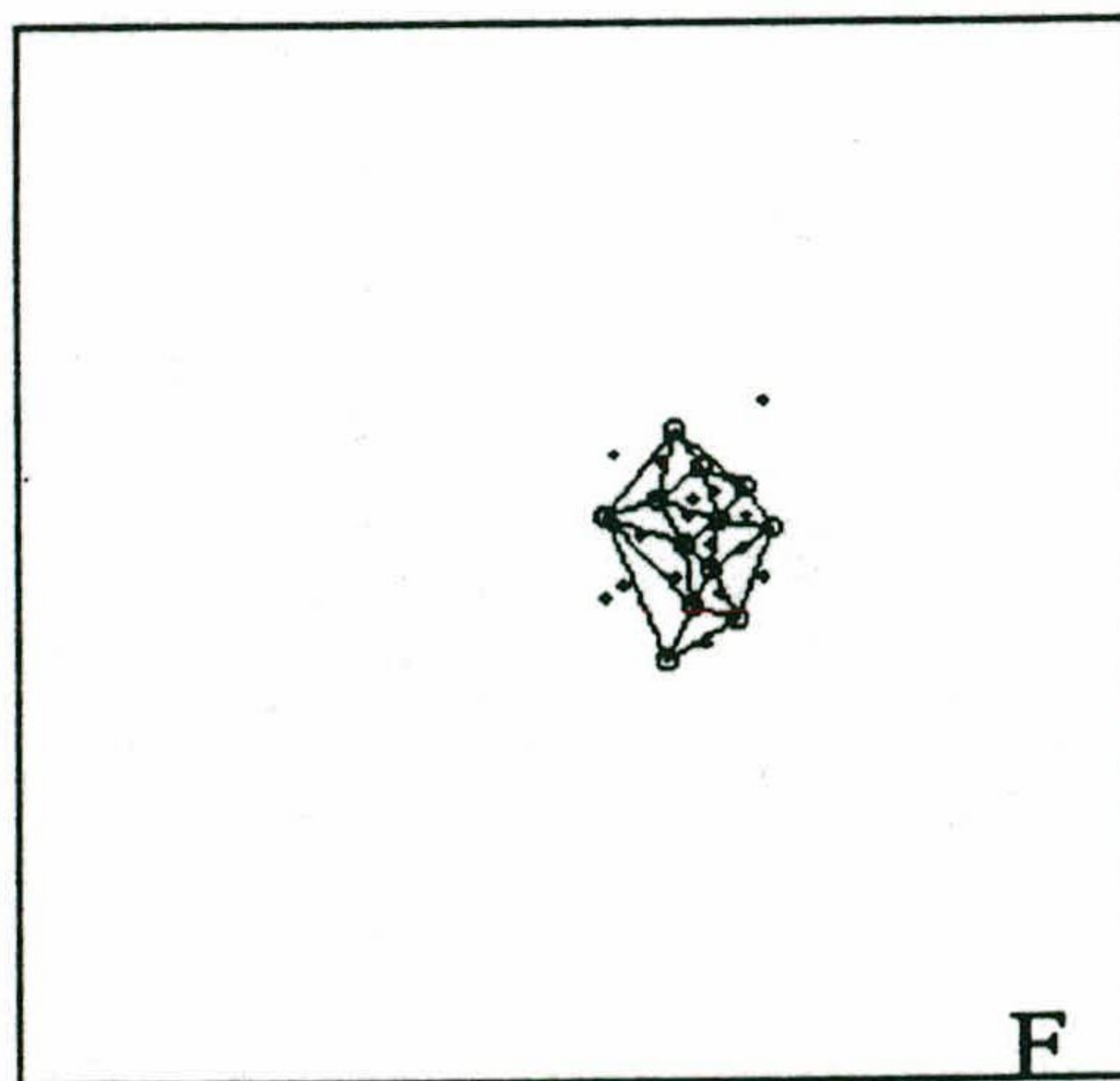
prb04124



D



prb03114



F



---

## II-4-3 DESCRIPCION DE LAS FUNCIONES DEL SOFTWARE ESPECIFICO DE ANALISIS AUTOMATICO DE LOS NUCLEOS Y GRUPOS GRANDES

### II-4-3-A Segmentación en caso de no solapamientos nucleares y segmentación de grupos grandes

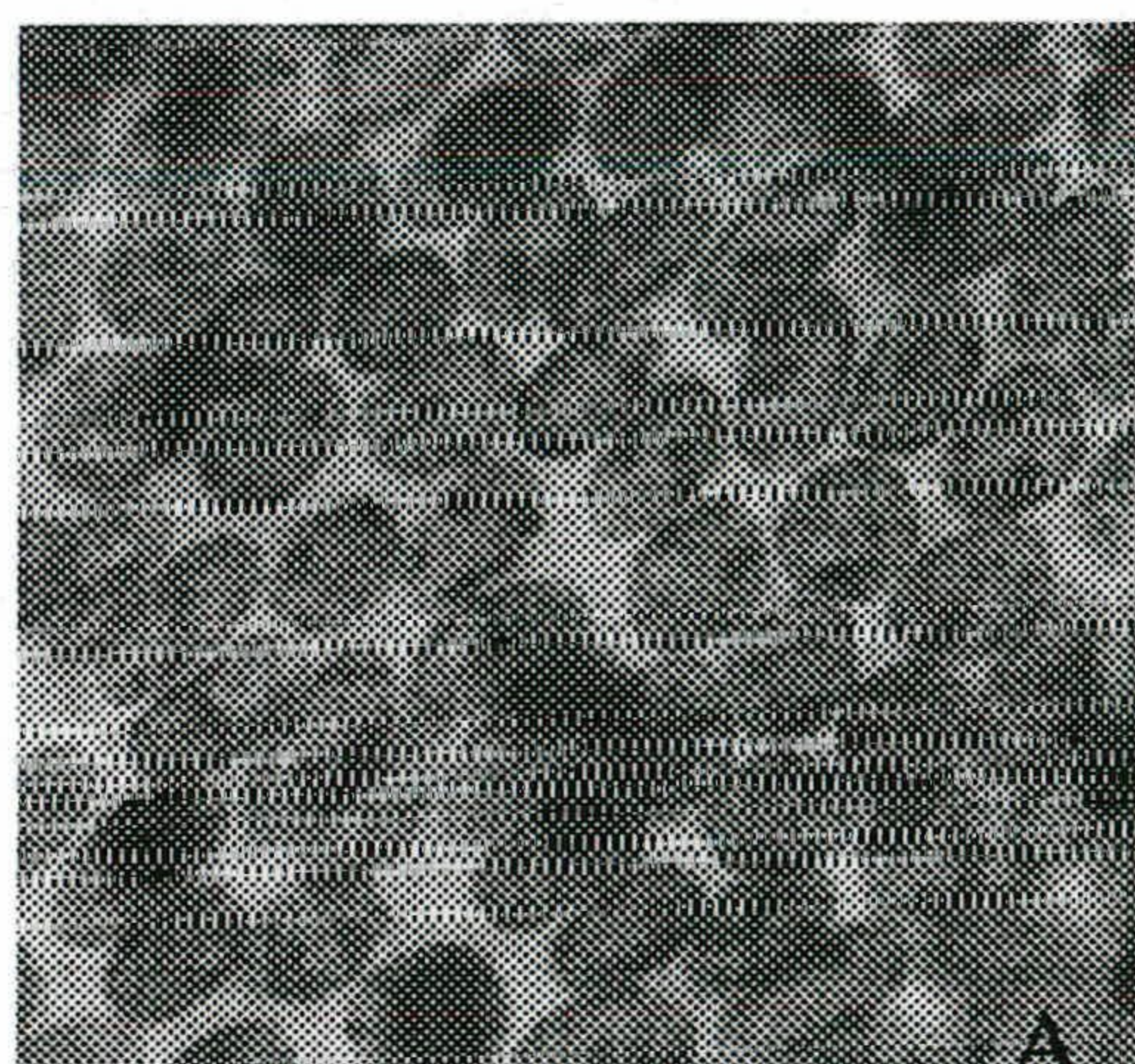
La primera aproximación a la segmentación de los núcleos celulares fué la de segmentar con la información suministrada por el histograma. En éste sentido son muchos los algoritmos que para hacer una umbralización o segmentación automática sobre una imagen digital, hacen uso de técnicas basadas en el histograma. Esto se debe a que, normalmente en una imagen, objetos distintos se corresponden con niveles de grises próximos en el histograma; en ése caso el problema de la segmentación automática de los núcleos queda reducido a la determinación de los umbrales sobre los que en el histograma se separan unos objetos de otros. Referente a ello cabe citar los trabajos de Luijendijk (1991) en el que se describe una técnica para la selección automática del umbral de segmentación basada en el recuento de regiones 4-conectadas (relación del pixel central con sus 4 vecinos más próximos) o el trabajo de Stoel (1990) en el que se diseñan algoritmos de segmentación sobre imágenes biomédicas mediante un proceso de equalización interactiva del histograma. En todos los casos, y de una forma u otra, se tratan de determinar propiedades que cumplen los objetos a segmentar sobre el histograma que los representa, de forma



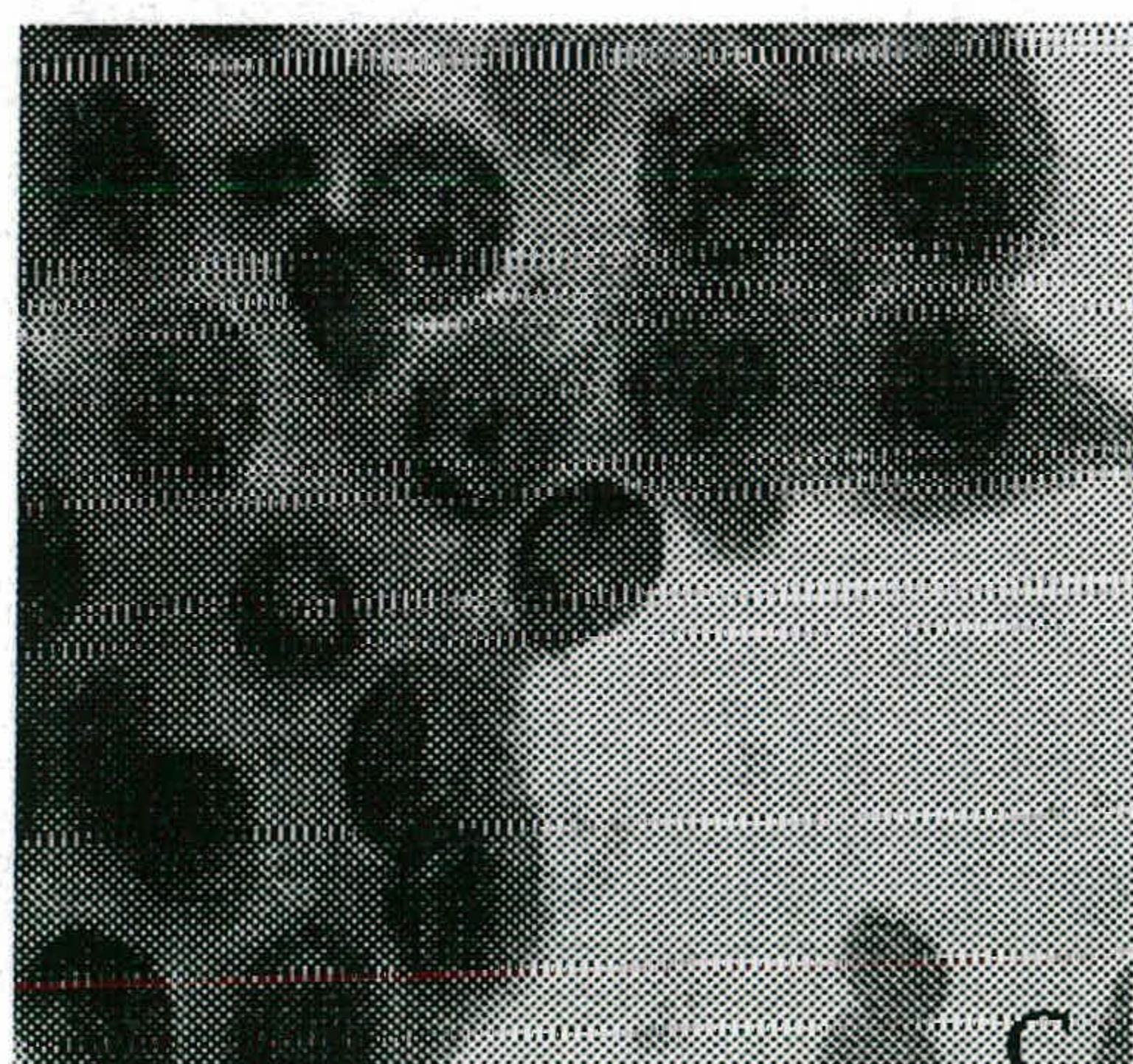
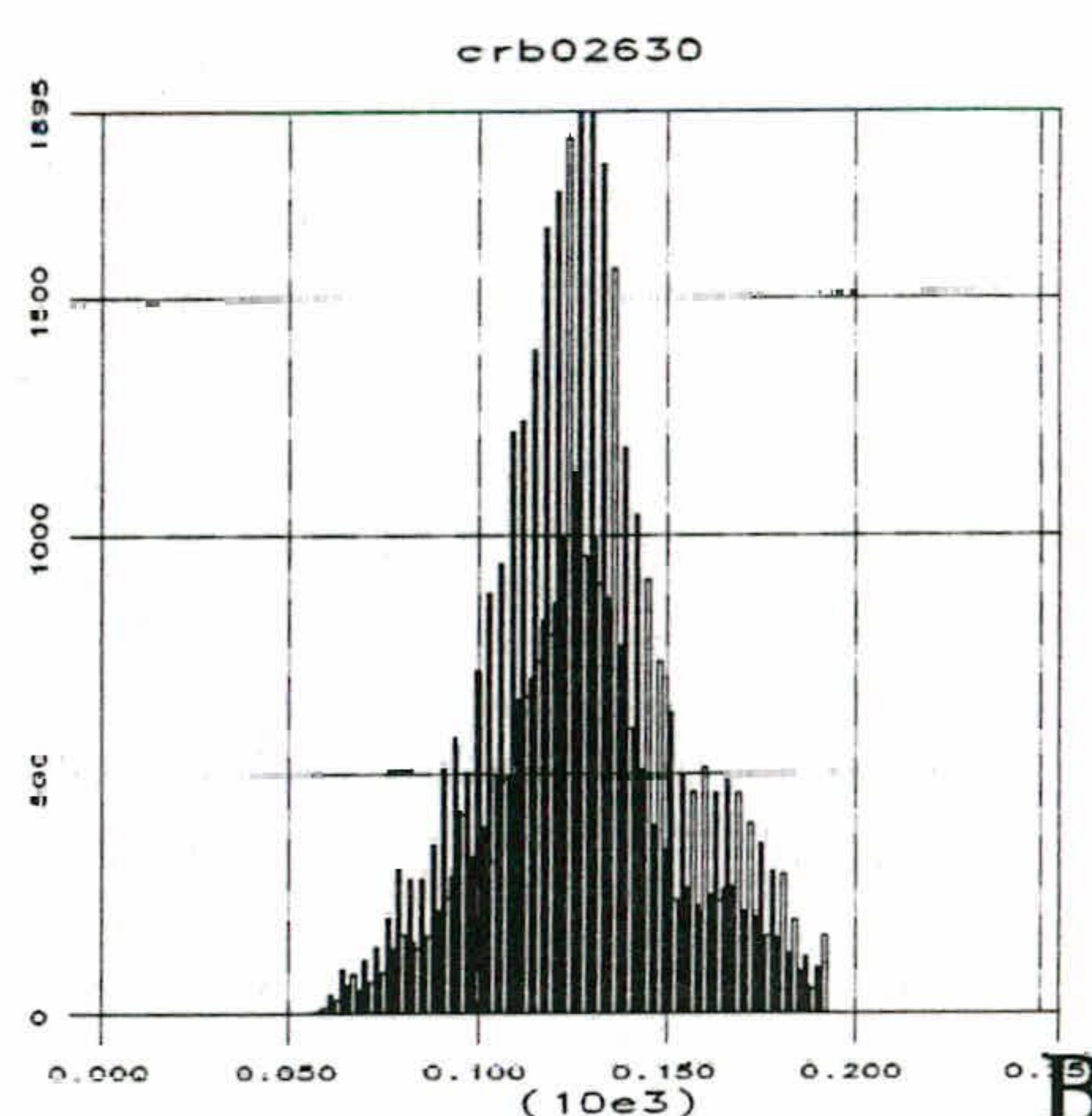
que sea posible aislarlos de forma automática.

A la vista de las formas que presentan los histogramas de las imágenes citológicas captadas (ver Fig.10) se pensó en determinar puntos críticos en la curva que daba el histograma: **determinación de máximos, mínimos y puntos de inflexión**, para ver si existía algún tipo de relación entre éstos puntos y las segmentaciones a que daban lugar.

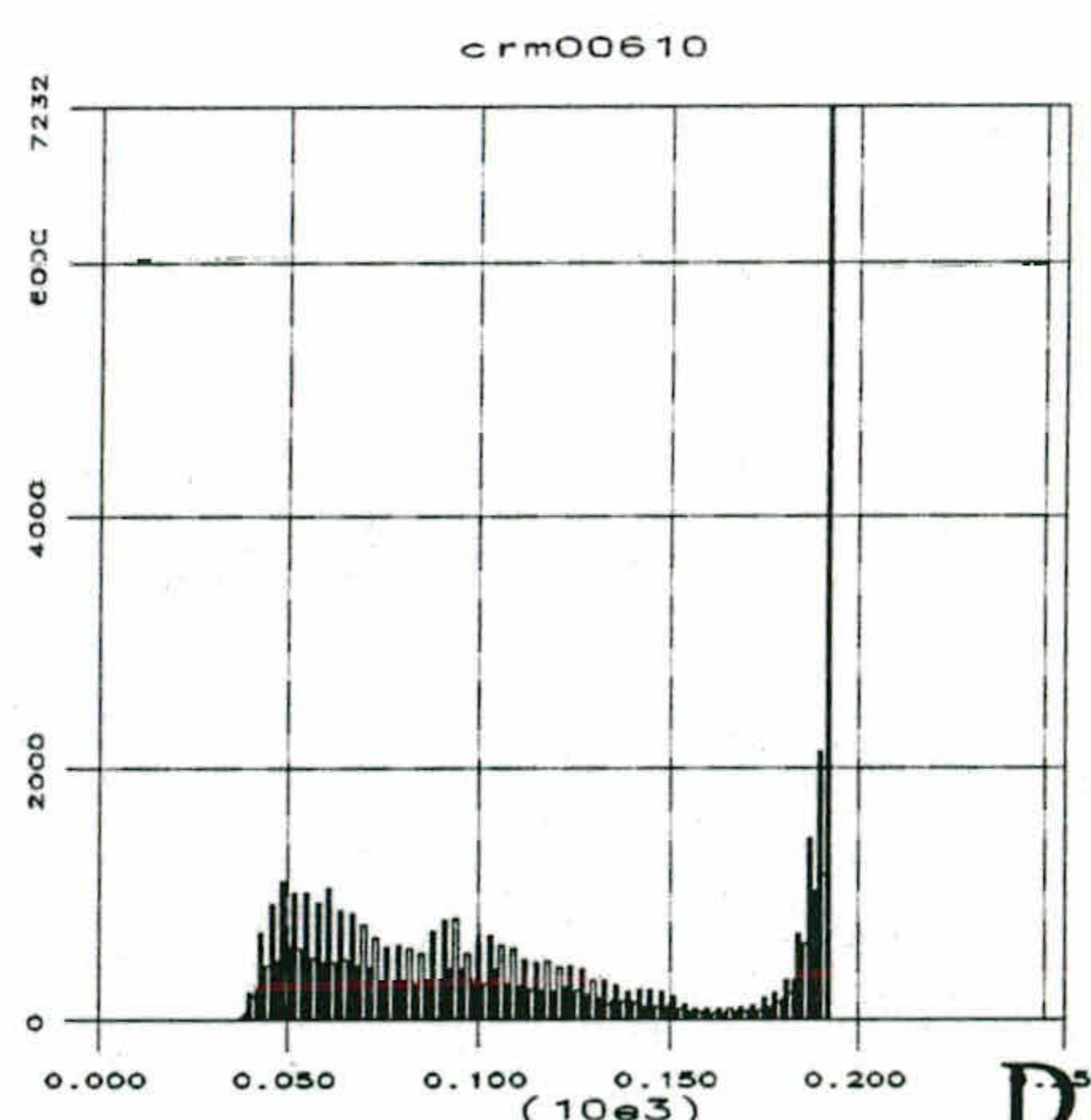
Figura 10: Histograma de las imágenes



crb02630



crm00610





---

En un primer momento se intentó determinar esos puntos críticos sobre el histograma acumulado. Sin embargo se abandonó esta línea debido a que en muchas imágenes las diferencias existentes en el crecimiento de esa curva eran tan pequeñas que se hacía difícil encontrar un método para la determinación de esos puntos críticos. Los resultados obtenidos sobre los histogramas de las imágenes sí permitieron determinar las propiedades que cumplían los umbrales de segmentación buscados.

#### ALGORITMO DESARROLLADO

El algoritmo de segmentación automática desarrollado tiene una parte común a los dos tipos de imágenes tratadas (núcleos de células y grupos grandes de células) y una parte específica para cada uno de esos tipos. A muy alto nivel de abstracción, el algoritmo quedaría descrito por:

- 1 Determinación de los puntos críticos sobre el histograma.
- 2 Agrupamiento de dichos puntos para eliminar candidatos.
- 3 Dependiendo del tipo de imagen a tratar y del objeto a segmentar, binarizar la imagen con alguna combinación de los puntos encontrados.

Se verá a continuación con más detalle cómo se ha abordado cada uno de esos puntos:

#### 1 DETERMINACIÓN DE PUNTOS CRÍTICOS

Para calcular máximos, mínimos y puntos de

---



---

inflexión se utilizan las propiedades analíticas sobre el signo que toman las derivadas de la función que determina el histograma en esos puntos críticos. Así pues necesitamos calcular al menos la primera y segunda derivada para encontrar los puntos. Debido al carácter discreto que tiene la función que representa al histograma de una imagen, el cálculo de la primera y la segunda derivada se resolvió utilizando la aproximación dada por Mokhtarian (1986) al cálculo discreto de las derivadas.

## 2 AGRUPAMIENTO DE PUNTOS CRÍTICOS

Una vez calculados los puntos críticos sobre el histograma se lleva a cabo un proceso de agrupamiento de los mismos debido a que el ruido de la función que representa el histograma puede dar lugar a fuertes oscilaciones en la tendencia de la curva, lo que puede generar distintos puntos críticos sobre una zona reducida de la curva cuando en realidad todos esos puntos podrían ser el mismo.

El rango o radio de acción del algoritmo de agrupamiento se da a través de un parámetro  $\epsilon$  que, en iteraciones pasadas sucesivas sobre los puntos críticos encontrados, los agrupa mediante la media dada por:

$$\text{Nuevo\_Punto} = \frac{\sum_{i=1}^{i=n} p(i)}{n}$$



---

donde:

- Nuevo\_Punto es el nuevo punto resultado del agrupamiento
- n es el número de puntos críticos pertenecientes a un grupo de radio  $\epsilon$ .
- p(i) es el valor de la abscisa de un punto crítico dado.

El algoritmo quedó pues como sigue (repetir hasta estabilizar valores):

- Dividir el histograma en trozos de tamaño  $\epsilon$
- Sea n el número de puntos críticos en un trozo cualquiera y p(i) la abscisa en esos i puntos.
- El punto que representa el trozo de tamaño  $\epsilon$  es:

$$\text{Nuevo\_Punto} = \frac{\sum_{i=1}^n p(i)}{n}$$

Un estudio experimental sobre la base de Datos de imágenes determinó que una elección razonable para el parámetro era  $\epsilon=3$ .

### 3 SEGMENTACIÓN A PARTIR DE LOS PUNTOS ENCONTRADOS

Notando P al conjunto de los puntos críticos encontrados sobre el histograma  $P = \{p^1, p_2, \dots, p_{n-1}, p_n\}$ . Tras un estudio experimental sobre la Base de Datos de imágenes, se determinó una elección razonable para los parámetros:

■ Núcleos en grupos. El umbral de segmentación seleccionado es:

$$P_1 + \frac{(P_2 - P_1)}{2}$$

■ Citoplasma de grupos. El umbral seleccionado es:

$$P_{n-1} + \frac{(P_n - P_{n-1})}{2}$$

■ Núcleos de células. Se obtienen los mejores resultados segmentando por:

$$P_{n-2}$$

Como postprocesamiento al proceso de segmentación se escogen las regiones resultantes en la imagen



---

correspondientes a componentes conexas representando células no solapadas y que no intersecten con los bordes de la imagen, rellenando, si es necesario, pixeles interiores que no hayan sido considerados por el algoritmo de segmentación (relleno de agujeros). El análisis de textura se efectúa sobre la zona de la imagen correspondiente a la componente célula del proceso de segmentación, es decir, al binarizar la imagen se tienen dos zonas: zona de células y zona de fondo, el análisis textural se hace sobre la primera. (Figura 11).

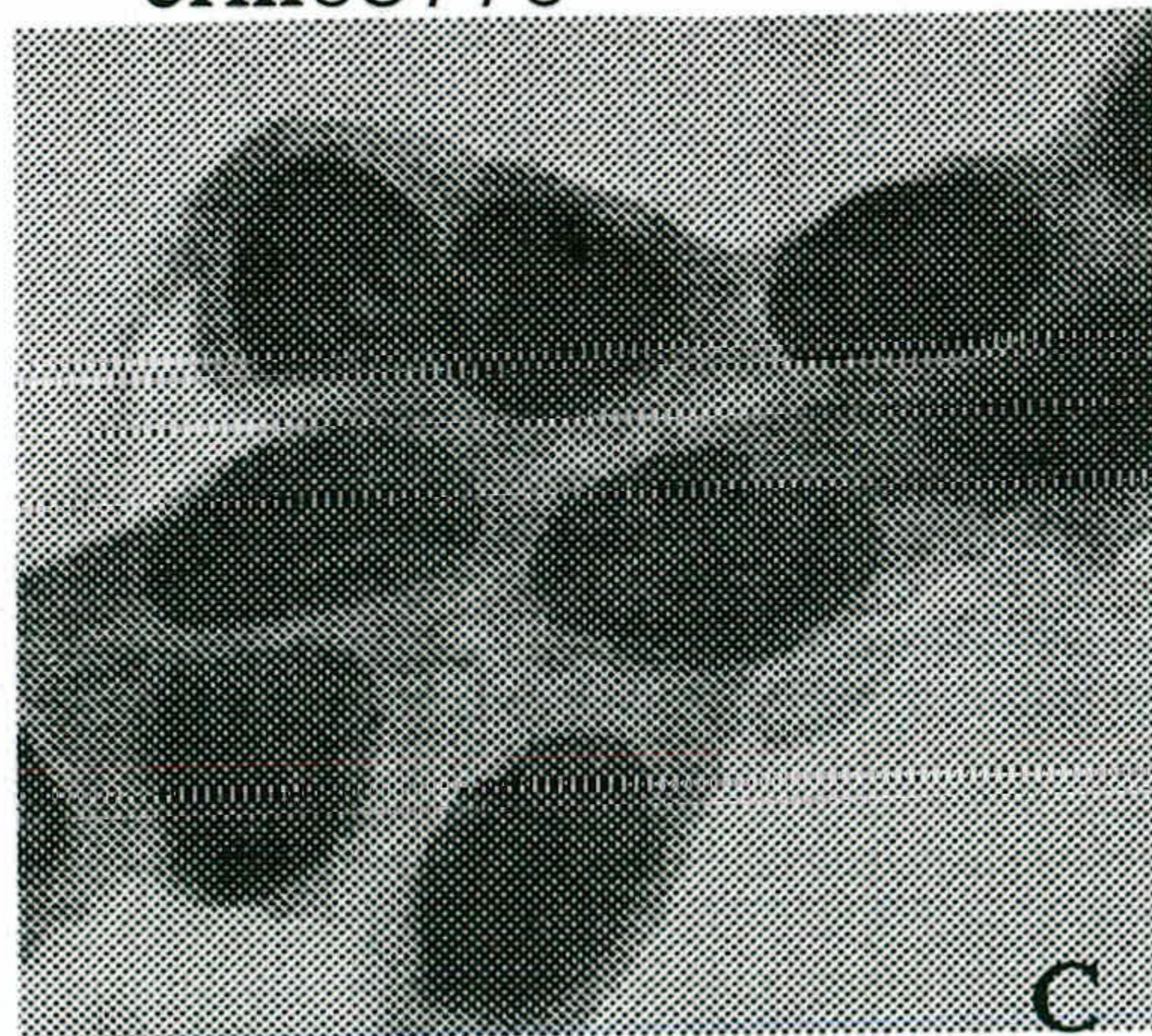
Figura 11: Ejemplos de segmentación para el caso de no solapamientos nucleares



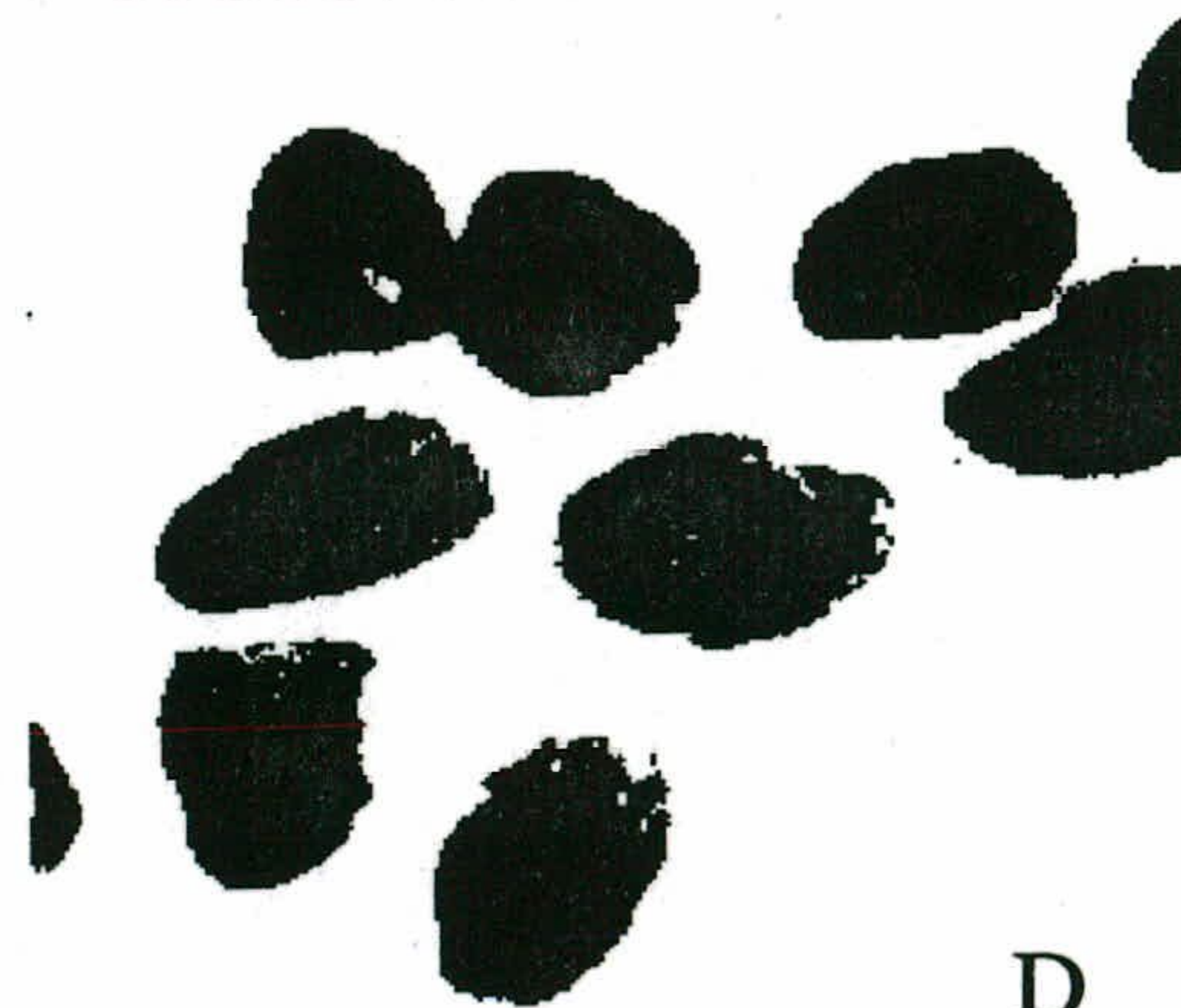
crm08770



cbb08770



crm08760



cbb08760

D



---

Los resultados con ésta aproximación eran muy prometedores en las imágenes en que no aparecían solapamientos pero cuando éstos estaban presentes, lo que era más frecuente en los casos malignos, la aproximación no resultaba totalmente correcta. Se optó, por tanto, en éstos casos, por refinar la aproximación para que segmentara la mayoría de los núcleos celulares, aún a pesar de los solapamientos.

#### II-4-3-B Segmentación de los nucleos para el caso de solapamientos

El desarrollo del refinamiento de la aproximación inicial llevada a cabo en el seno del grupo de Análisis de Imágenes del Depto. de Ciencias de la Computación (Pérez de la Blanca 1993 y Garrido 1994) tuvo como origen, como antes indicábamos, el tratar con los problemas que planteaban los solapamientos y las severas condiciones de ruido que aparecían en la mayoría de las imágenes que se procesaron. El fracaso en éstos casos de los métodos de segmentación basados en el histograma hizo que se diseñara otra metodología que hacía uso de las informaciones que aportaba los estimadores de frontera (Canny 1986) desde distintas escalas informacionales con el objeto de desarrollar un método de preprocesamiento de las cadenas que definieran las fronteras. La estimación final de las formas se hace a través de un modelo de patrón deformable (de aproximación a una



---

elipse) en el que se hace intervenir como información a priori, la obtenida en la etapa anterior de procesamiento de las cadenas de fronteras. En el presente trabajo se ha hecho uso de detectores multiresolución (su funcionamiento depende de un parámetro que no tiene valor fijo), desarrollados a partir de los trabajos de Marr (1980) en los que los cambios de intensidad se analizan a distintas escalas (dependiendo del valor del parámetro), combinando posteriormente los resultados para obtener la imagen de fronteras final. El modelo desarrollado es invariante frente a traslaciones, rotaciones, cambios de escala, deformaciones locales y ruido (Garrido 1994 y 1995).

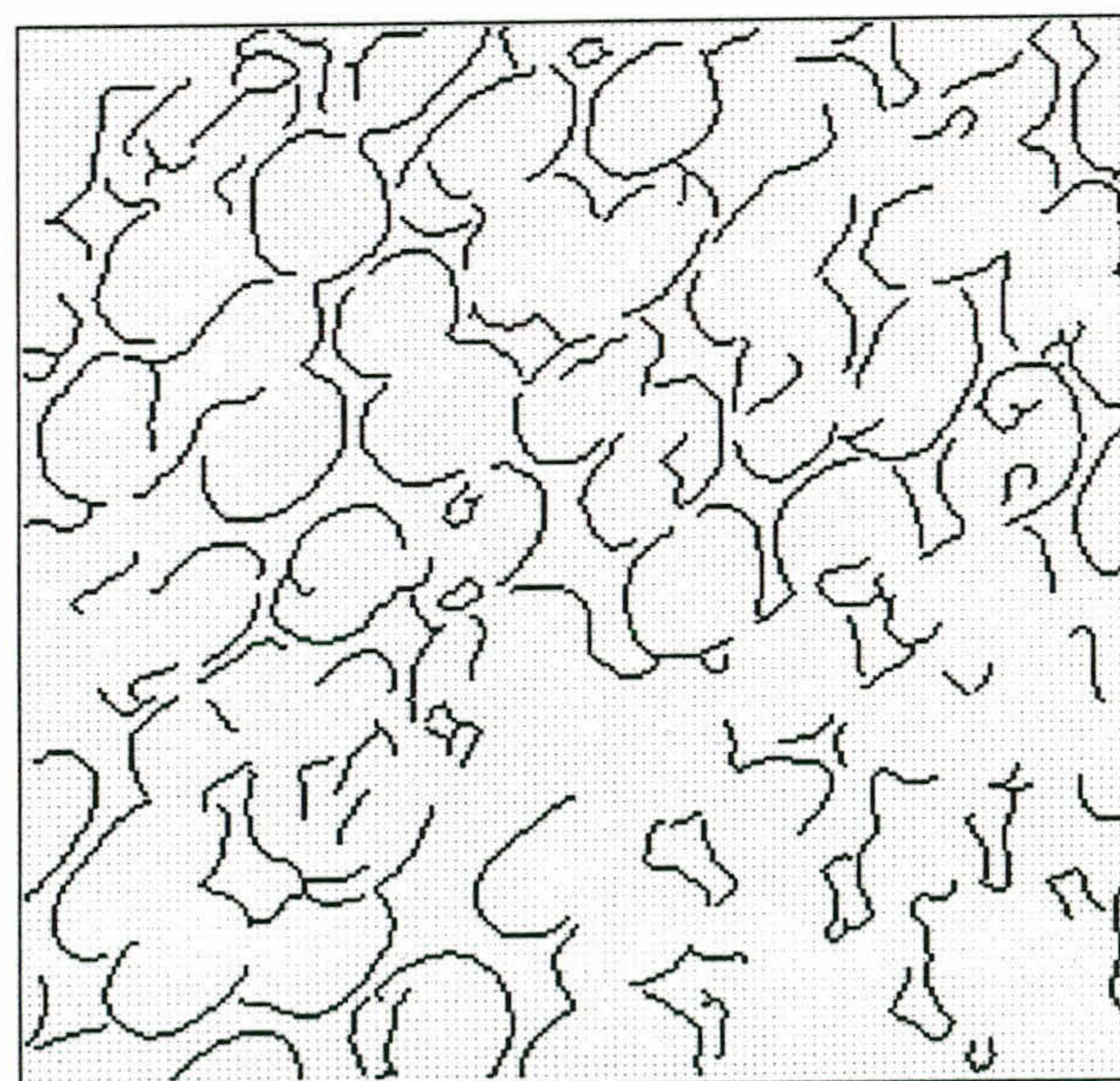
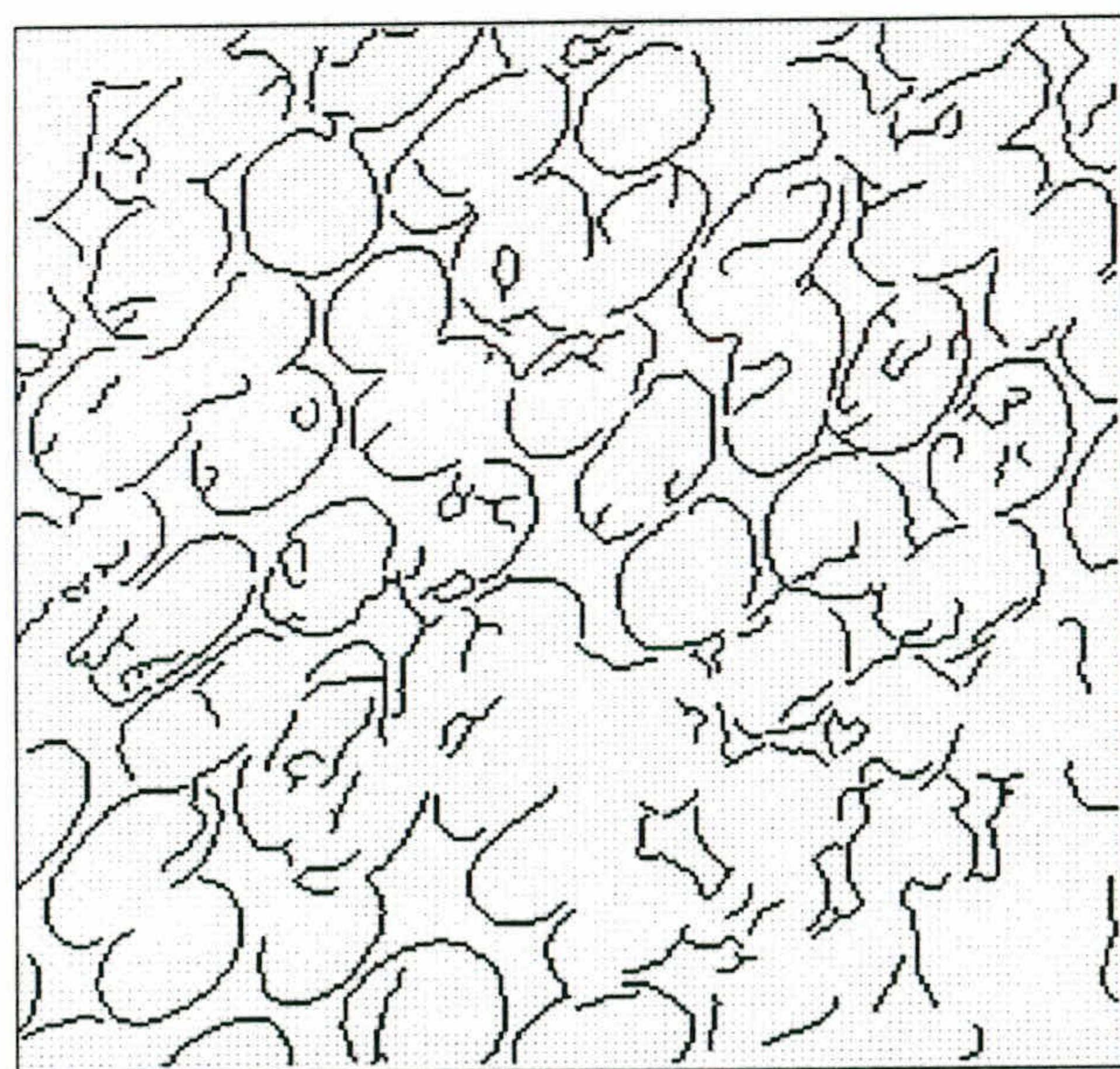
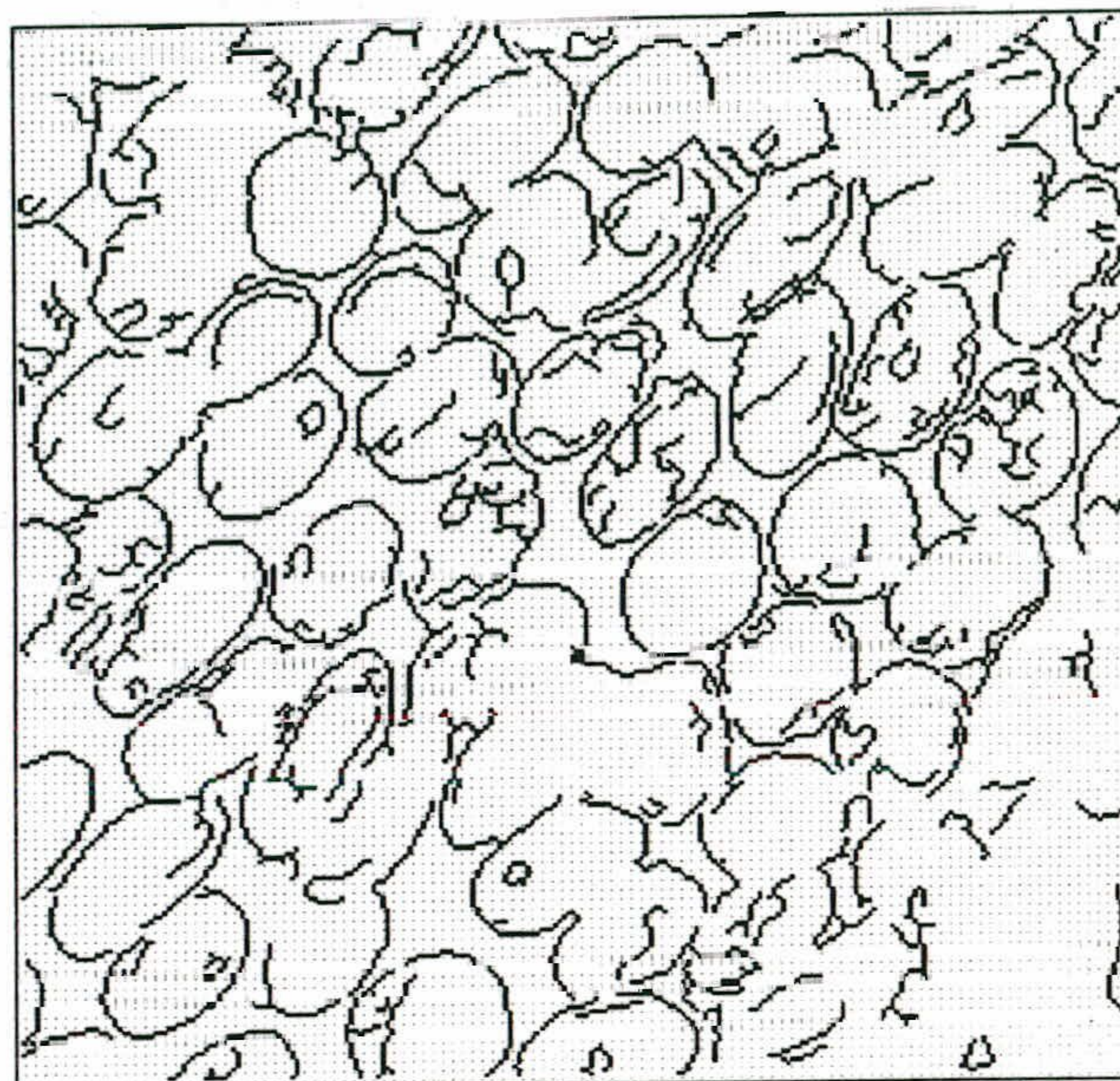
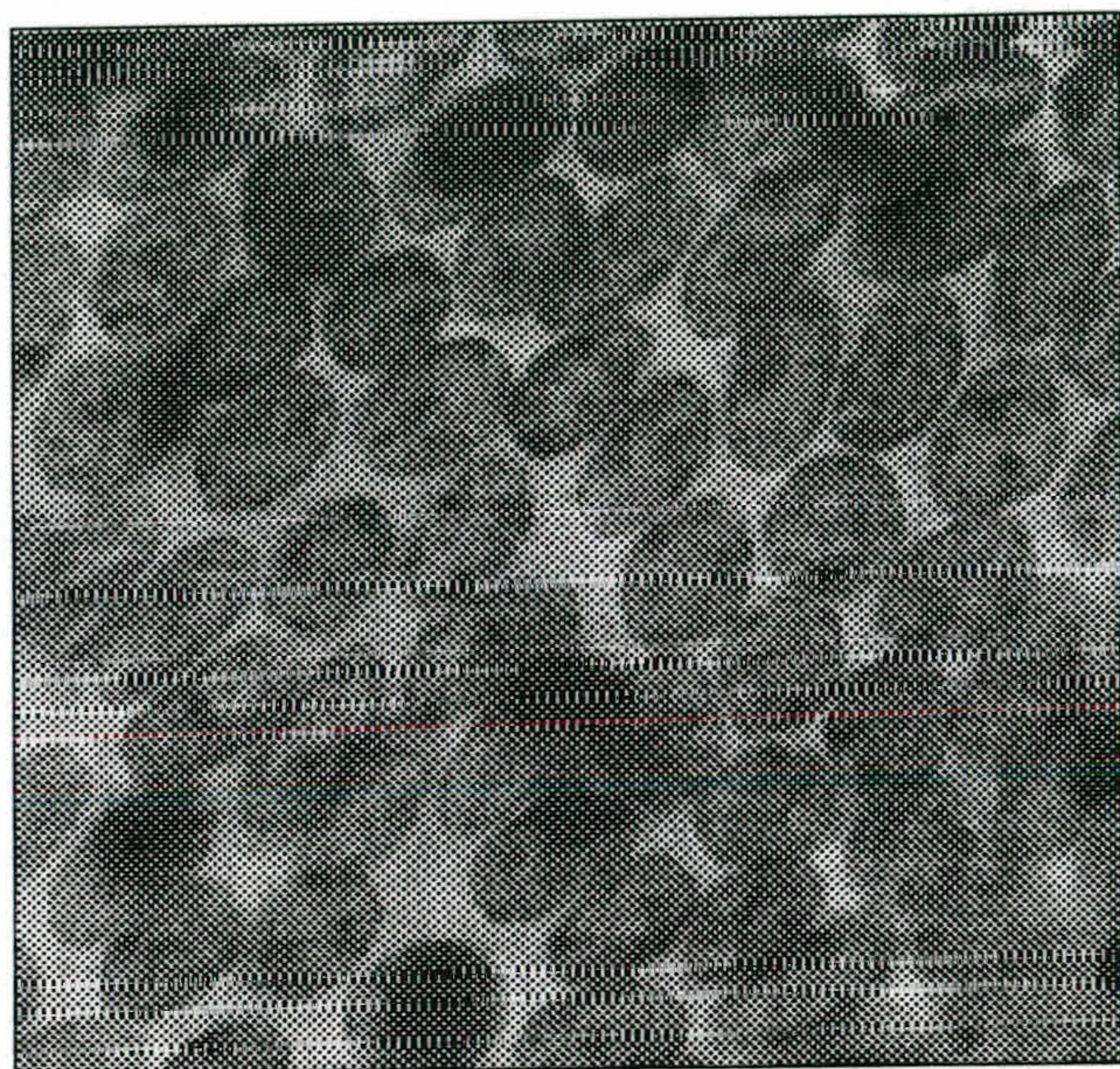
Los pasos del algoritmo implementado son los siguientes:

- 1 Aplicación de un algoritmo de detección de fronteras (Canny 1986) a distintas resoluciones, y para cada una de las imágenes de fronteras:
  - 1.1 Busca, corta y orienta las cadenas que delimitan las fronteras de las células.
  - 1.2 Une las cadenas hasta obtener contornos cerrados.
- 2 Integra la información de las distintas resoluciones:
  - 2.1 Aproxima la célula mediante una elipse
  - 2.2 Hace la partición final de las células.

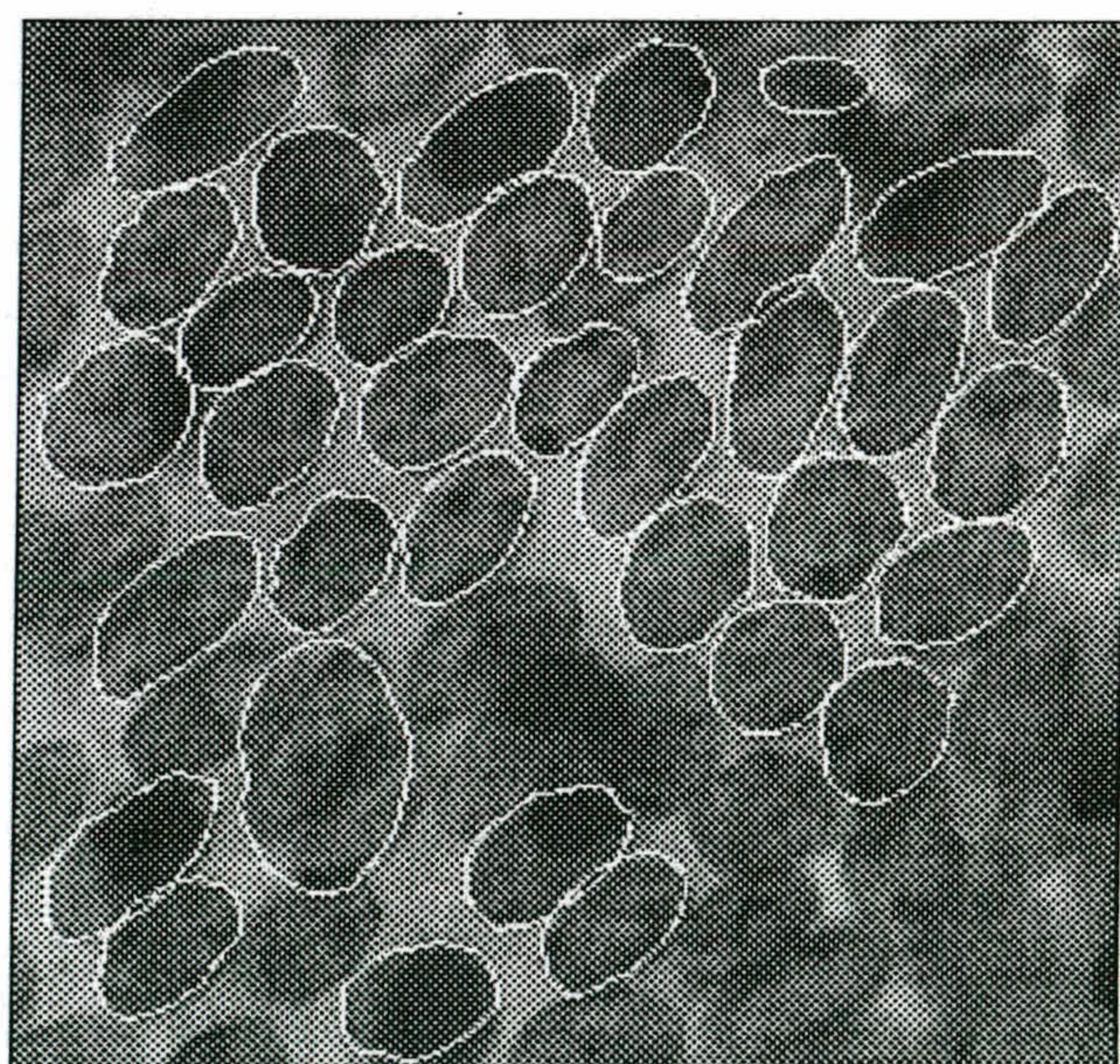
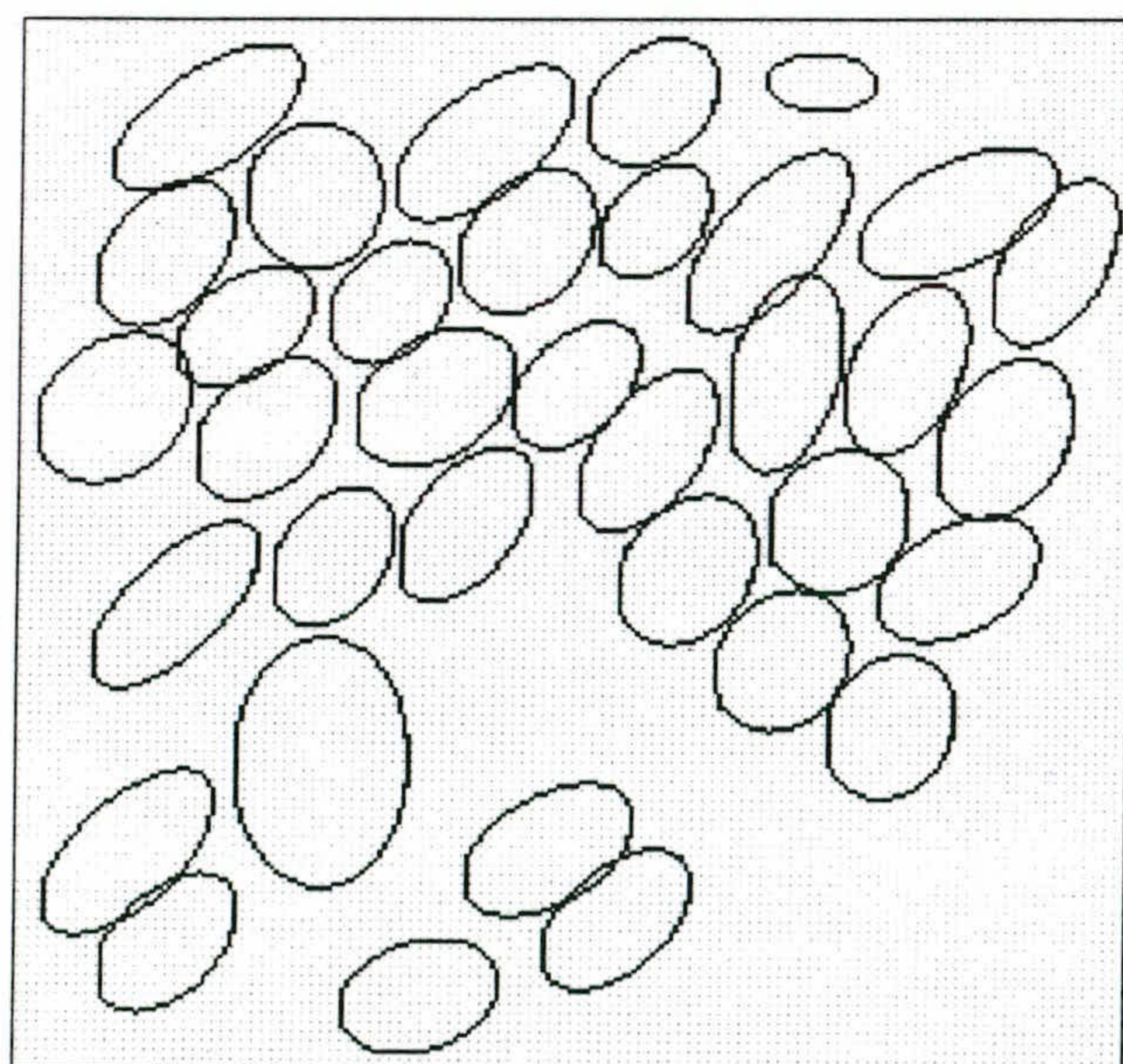
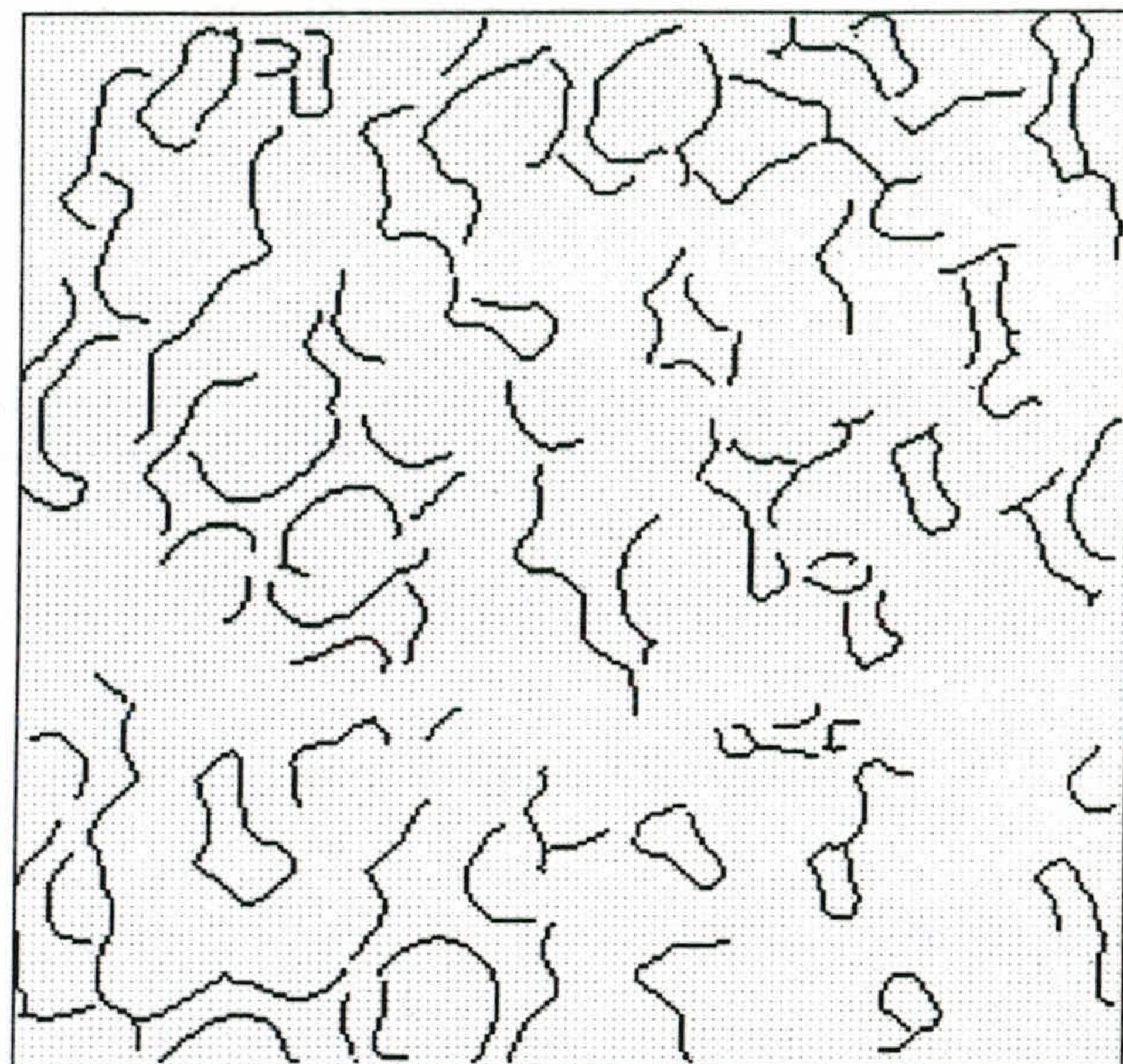
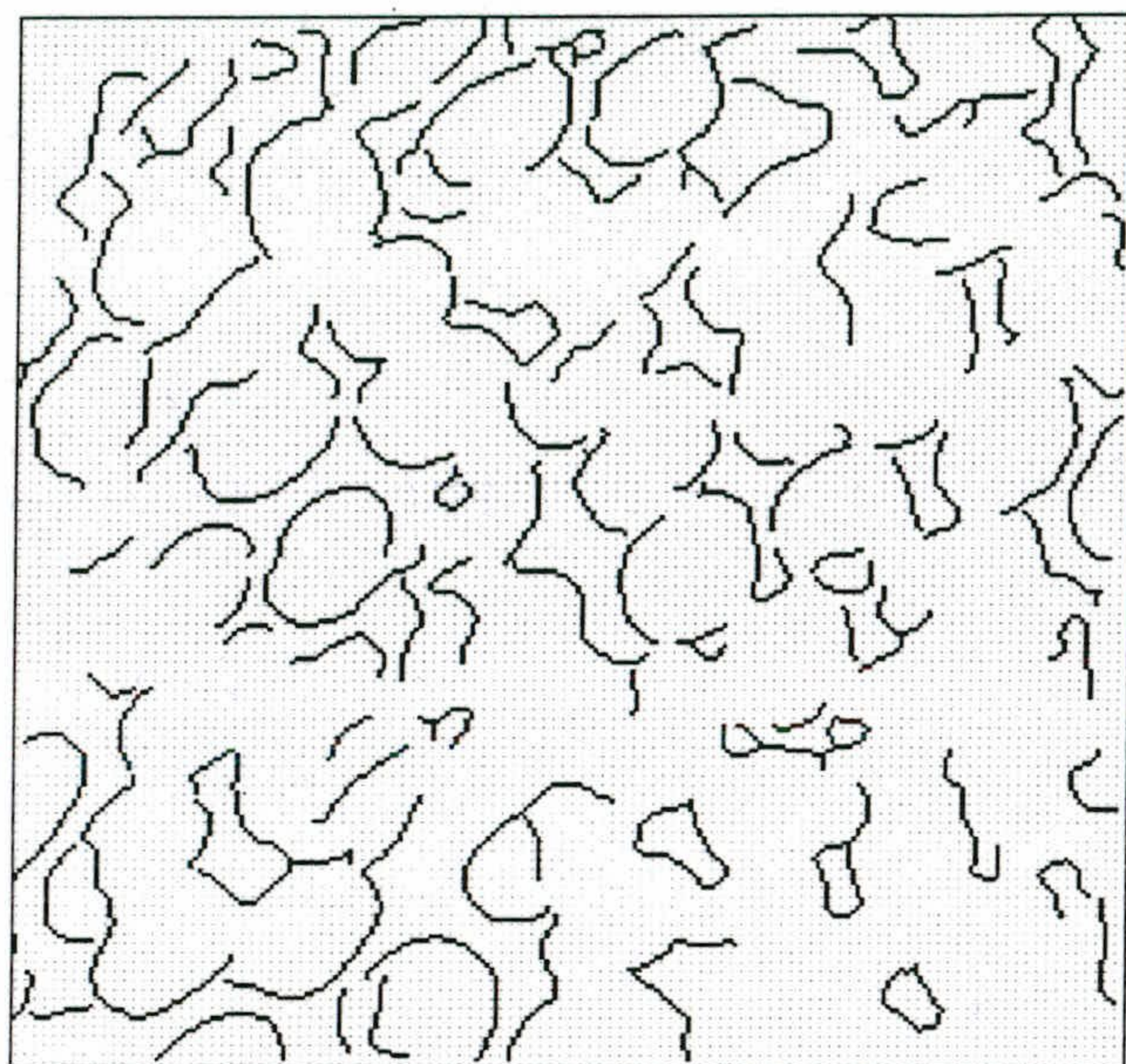
Ejemplo de aplicación de la aproximación en presencia de solapamientos puede verse en la Figura 12.



Figura 12









---

## II-5 EL ANALISIS DE LOS DATOS

El análisis de los datos se va a realizar con una doble metodología : **Análisis Discriminante y metodología CART** (árboles binarios de clasificación y regresión), para luego hacer una discusión comparativa de los resultados. Con ambos métodos se pretenden construir reglas de predicción que nos asignen cada caso o imagen analizados a una clase: **BENIGNO** o **MALIGNO**.

Es bien conocido el funcionamiento y la metodología del Análisis Discriminante, en cuanto a CART, indicar que funciona seleccionando las reglas de predicción lo más exactas posibles usando la mínima cantidad de variables (Aus 1986, Feernández Valdivia 1991). La construcción del árbol se realiza sobre un "*conjunto de aprendizaje de datos*", para ello Cart examina cada una de las variables de la forma  $X_i < C$  (siendo C una constante que está entre el mínimo y el máximo valor de  $X_i$ ), eligiendo las variables y sus valores C para cada división o nodo del árbol que mejor separe las clases. El tamaño adecuado del árbol será aquel que minimice el error real, estimando éste utilizando conjuntos test, escogiendo para testear usualmente el 10 % de los casos, Cart selecciona el menor árbol que tiene la menor tasa de error estimado.

Con las Matrices de Clasificación del resultado de los test, en uno u otro caso, obtenemos la *Especiticidad*,



---

*Sensibilidad, Eficiencia, Valor Predictivo Positivo y Valor Predictivo Negativo del método.*

## II-5-1 CLASIFICACION

La clasificación benignas-malignas se realizó en base a variables nucleares y de los grupos, ciñéndonos a aquellas características que el experto tiene en cuenta en su diagnóstico y que podían ser cuantificables y tratables, lo más automáticamente posible, con técnicas de análisis de imagen. Para resolver en los grupos la variable AN/AC se aplicó la macro desarrollada en el programa Visilog con la que se obtuvieron dos índices de relación. El software desarrollado por el Grupo de Análisis de Imagen del Departamento de Ciencias de la Computación de la Universidad de Granada nos permitió resolver el grueso del problema, ordenando las variables en tres grupos según sus características con el objeto de determinar qué peso específico tenía cada uno de ellos en la discriminación:

CLASIFICACIÓN EN BASE A VARIABLES PLANIMÉTRICAS DE LOS NÚCLEOS, EN BASE A TEXTURAS DE LOS NÚCLEOS Y EN BASE A DISTANCIAS ENTRE ELLOS DENTRO DEL GRUPO.



---

## II-5-1-A Clasificación en base a la Planimetría Nuclear

Los descriptores simples medidos son:

**Area (A):** Número de píxeles contenidos en la región R

**Perímetro (P):** Número de píxeles de la frontera de la región.

**Feret de las x (Fx) y de las y (Fy)** (Diámetro mayor y menor)

**Estericidad (E)** =  $(1/4\pi) * (P^2/A)$

**Elipticidad: El** =  $Fy/Fx$

Las medidas que se consideran para cada imagen son la media y desviación y, eventualmente, el valor máximo. Esjtos mismos descriptores fueron utilizados para los grupos grandes.

## II-5-1-B Clasificación en base a Texturas

Nosotros, por el tipo de textura de que disponemos, hemos utilizado para la clasificación la aproximación estadística, dejando a un lado la estructural por la dificultad que supone en nuestras imágenes la búsqueda e identificación de primitivas, su modelización y las dificultades computacionales y de estimación para un problema que, en principio, parece resoluble por otras técnicas.

El problema fundamental que se plantea es el de una selección inicial del conjunto de medidas estadísticas,



---

algunas de ellas sin un significado intuitivo obvio, y la interpretación final de los resultados del proceso discriminante. Por tanto la idea para el proceso de clasificación es medir, para cada una de las clases, un conjunto de medidas que en bloque formarán un vector y a través de éste, y usando técnicas estadísticas, construir el clasificador (con las medidas que hayan resultado más significativas) que posteriormente podrá usarse en la clasificación de nuevas observaciones.

El conjunto de medidas que inicialmente se consideraron son los que aparecen en los distintos trabajos sobre el tema (Galloway 1974, Haralick 1979, Levine 1980...etc), algunas de ellas motivadas por estudios del sistema visual humano que muestran que los estadísticos de segundo orden basados en las medidas obtenidas por las matrices de dependencia espacial de los niveles de gris, juegan un papel importante en la discriminación de las texturas. Primeramente se realizó experiencias pilotos con estas medidas para determinar cuales de ellas daban buenos resultados con nuestras imágenes.

Esquemáticamente las medidas que finalmente se contemplaron en éste trabajo pueden resumirse en la siguiente lista:



---

A-MEDIDAS BASADAS EN EL HISTOGRAMA:

Sea  $H(l_k)$  el valor del histograma normalizado para un nivel de gris  $l_k$  ( $k=1\dots255$ ). Las medidas que consideramos a partir de  $H$  son sus 4 primeros momentos centrales y sus correspondientes desviaciones estándar (SD): el momento de orden 1 ( $\mu$ ) es la media de los niveles de gris, el momento de orden dos ( $\sigma$ ) es la varianza de los niveles de gris (ambos orientan sobre la uniformidad o regularidad de una región), el momento de orden tres que orienta sobre la simetría ( $\mu_3$ ) y el momento de orden 4  $\mu_4$ .

$$\mu = \sum_{k=1}^{255} l_k * H(l_k). \text{ (Momento de orden 1). Media de los niv. de gris.}$$

$$\sigma = \sum_{k=1}^{255} (l_k - \mu)^2 * H(l_k). \text{ (Momento de orden 2). Varianza ,, ,,}$$

$$\mu_3 = (1/\sigma) * \sum_{k=1}^{255} (l_k - \mu)^3 * H(l_k). \text{ (Momento de orden 3).}$$

$$\mu_4 = (1/\sigma^4) * \sum_{k=1}^{255} (l_k - \mu)^4 * H(l_k). \text{ (Momento de orden 4).}$$

B-MEDIDAS BASADAS EN DIFERENCIAS DE NIVELES DE GRIS

En un histograma 2-D se representa el número de ocurrencias en la imagen de cada diferencia de niveles de gris según un determinado vector de desplazamiento. Llamamos  $H(d_s, \delta)$  a la probabilidad de que una diferencia dada  $d_s$  ocurra según el vector de desplazamiento  $\delta$ . Las medidas que consideramos a partir de  $H$  son las siguientes:



---


$$\mu_t = \sum_{k=1}^{I=1} d_k * H(d_k, \delta). \text{ Media de las diferencias.}$$

$$C_t = \sum_{k=1}^{I=1} d_k^2 * H(d_k, \delta). \text{ Medida de contraste.}$$

$$E_t = - \sum_{k=1}^{I=1} H(d_k, \delta) * \log H(d_k, \delta) \text{ Medida de uniformidad.}$$

$$A_t = \sum_{k=1}^{I=1} H(d_k, \delta)^2. \text{ Medida de contraste.}$$

#### C-MEDIDAS BASADAS EN COOCURRENCIAS DE NIVELES DE GRIS

Se considera el histograma 2-D del número de veces que parejas de niveles de gris ocurren en una relación espacial dada. Su valor depende de la fiabilidad con que se captura la textura. Sea  $H(l_r, l_s, \delta)$  la probabilidad de que dos píxeles situados uno respecto del otro según el vector de posición  $\delta$ , ocurran uno con nivel de gris  $l_r$  y otro con nivel de gris  $l_s$ . Siendo la constante de normalización en éste caso la suma de los valores de la matriz de coocurrencias. Las medidas consideradas son:

$$CC_1 = \sum_{r=1}^{I=1} \sum_{s=1}^{S=1} (l_r - l_s)^2 * H(l_r, l_s, \delta). \text{ Medida de periodicidad.}$$

$$CC_2 = \sum_{r=1}^{I=1} \sum_{s=1}^{S=1} [1/(1+(l_r - l_s)^2)] * H(l_r, l_s, \delta). \text{ Medida de homogeneidad local.}$$

$$CC_3 = \sum_{r=1}^{I=1} \sum_{s=1}^{S=1} H(l_r, l_s, \delta)^2. \text{ Medida de contraste.}$$

$$CC_4 = \sum_{r=1}^{I=1} \sum_{s=1}^{S=1} H(l_r, l_s, \delta) * \log H(l_r, l_s, \delta). \text{ Medida de uniformidad.}$$

$$CC_5 = [1/(\sigma_{l_r} * \sigma_{l_s})] \sum_{r=1}^{I=1} \sum_{s=1}^{S=1} (l_r - \mu_{l_r}) * (l_s - \mu_{l_s}) * H(l_r, l_s, \delta). \text{ Medida de correlación.}$$

Siendo  $\sigma_{l_r}$ ,  $\mu_{l_r}$  la desviación y la media del vector de probabilidad marginal obtenido sumando sobre las filas de  $H$  (análogo  $\sigma_{l_r}$ ,  $\mu_{l_r}$  sobre las columnas).



---

## II-5-1-C Clasificación en base a distancias nucleares en los grupos: Creación de Grafos Planares

Con los grafos planares se mide el desorden de los núcleos dentro del grupo, cuantificandoo de ésta forma una de las variables más importantes que el experto tiene en cuenta para discriminar benignidad de malignidad.

El grafo planar conecta cada centro nuclear o nodo con sus vecinos más próximos, con ello se determina las distancias entre núcleos y las zonas de influencia de unos sobre otros a través de 4 variables: la media y la desviación de las distancias y el máximo y mínimo de las mismas. El algoritmo construido utiliza las propiedades existentes entre los nodos del grafo y los polígonos de Voronoi para obtener un grafo a partir de triángulos que cumplen las condiciones de ser lo más regulares posibles.



---

II-5-1-D Clasificación en base a la relación AN/AC en los  
grupos

---

La Clasificación se basó en la obtención de dos índices: I e I'.

$$I = (AG-AC)/AG$$

$$I' = (AG-AC)/AC$$

Siendo AG = Area total del grupo

AC = Area del citoplasma

Por tanto AG-AC es el área total que ocupan los núcleos que se toman en el plano de enfoque dentro del grupo.



---

## RESULTADOS



---

### III RESULTADOS

Los rasgos cuantificables por la metodología elegida y para el objetivo propuesto quedaron divididos en tres grupos: DE GRUPOS GRANDES, DE GRUPOS PEQUEÑOS Y DE NUCLEOS. Los rasgos de los grupos grandes: Máximo del tamaño de los grupos de cada clase y forma de éstos no dieron información discriminadora en los resultados preliminares por lo que esta línea se abandonó. En los grupos pequeños, el clasificador utilizando índices de relación de áreas, nuclear y citoplasmática, clasificaba correctamente todos los casos malignos pero todos los benignos también los clasificaba como malignos. Los rasgos del extendido celular: presencia de necrosis tumoral, inflamación, presencia de estromas y presencia de células o núcleos sueltos semejantes a los integrados en los grupos, no eran fácilmente cuantificables con análisis de imagen o requerían arduo procedimiento interactivo. Los nucleolos tampoco se tuvieron en consideración por la imposibilidad de distinguirlos de grandes acúmulos cromatínicos a causa de la falta de especificidad de la tinción utilizada. El resultado del resto de las variables se expone a continuación.







---

De los 47 casos benignos (B) y de los 95 casos malignos (M) se obtuvieron los siguientes resultados:

MATRIZ DE CLASIFICACIÓN DE LA MUESTRA DE APRENDIZAJE

	Clase Verdadera	
	B	M
B	41	4
M	6	91

MATRIZ DE PROBABILIDAD DE LA MUESTRA DE APRENDIZAJE

	Clase Verdadera	
	B	M
B	0,87	0,04
M	0,13	0,96

MATRIZ DE CLASIFICACIÓN DE LA VALIDACIÓN CRUZADA

	Clase Verdadera	
	B	M
B	40	5
M	7	90

MATRIZ DE PROBABILIDAD DE LA VALIDACIÓN CRUZADA

	Clase Verdadera	
	B	M
B	0,85	0,05
M	0,15	0,95

Aplicando las fórmulas de la Eficacia (Ef), Sensibilidad (S), Especificidad (Es), Valor predictivo (VPP) positivo y valor predictivo negativo (VPN) de los datos obtenidos en la Matriz de Clasificación de la Validación Cruzada, obtenemos para éste experimento con v. planimétricas:



---


$$\begin{aligned}
 Et &= 100 * (vp + vn) / total = 91.5 \% \\
 S &= 100 * (vp / (vp + fn)) = 94.7 \% \\
 Es &= 100 * vn / (vn + fp) = 85.1 \% \\
 VPP &= 100 * vp / (vp + fp) = 92.8 \% \\
 VPN &= 100 * vn / (vn + fn) = 88.9 \%
 \end{aligned}$$

Siendo:

vp = verdaderos positivos = 90  
 vn = verdaderos negativos = 40  
 fp = falsos positivos = 7  
 fn = falsos negativos = 5

La importancia de las variables en orden decreciente, asignándole un valor de 100 a la más relevante, va en éste orden:

desviación del área	100
,, estericidad	98
,, feret y	89
,, feret x	79
,, perímetro	78
Feret X	76
perímetro	73
estericidad	72
area	72
Feret y	66
varianza niveles gris	58
circularidad	53
des. var. nivel gris	34
des. circularidad	19
elipticidad	15
des. media nivel gris	15
des. elipticidad	11
des. simetría	8
simetría	8
des. momento 4	5
media nivel de gris	4
momento 4	4



---

El orden de las variables, al añadirle los máximos de las 6 primeras queda como sigue:

desviación del área	100
,, esfericidad	98
máximo del perímetro	97
,, feret X	90
desviación feret Y	89
máximo área	89
,, esfericidad	88
,, feret Y	85
desviación feret x	79
,, perímetro	78
feret X	74
esfericidad	71
área	71
perímetro	71
feret Y	65
varianza niveles gris	58
circularidad	53
desv.varianza niveles	34
,, circularidad	22
máximo circularidad	19
elipticidad	15
desv. media niveles gris	15
,, elipticidad	11
,, simetría	8
simetría	8
desviación momento 4	5
momento 4	4
media de los niveles gris	4

En el primer grupo observamos que tienen más importancia las desviaciones de las variables que sus medias, esto se debe a que la media compensa los valores extremos. Estos valores extremos se debe a la heterogeneidad de las poblaciones celulares de los casos malignos, las desviaciones resalta esa diferencia frente a la mayor homogeneidad de los casos benignos. En el segundo grupo, al que se le ha añadido los máximos de las 6 primeras variables, parecen tener también bastante importancia diagnóstica, aunque las dos primera posiciones se mantienen y las matrices de

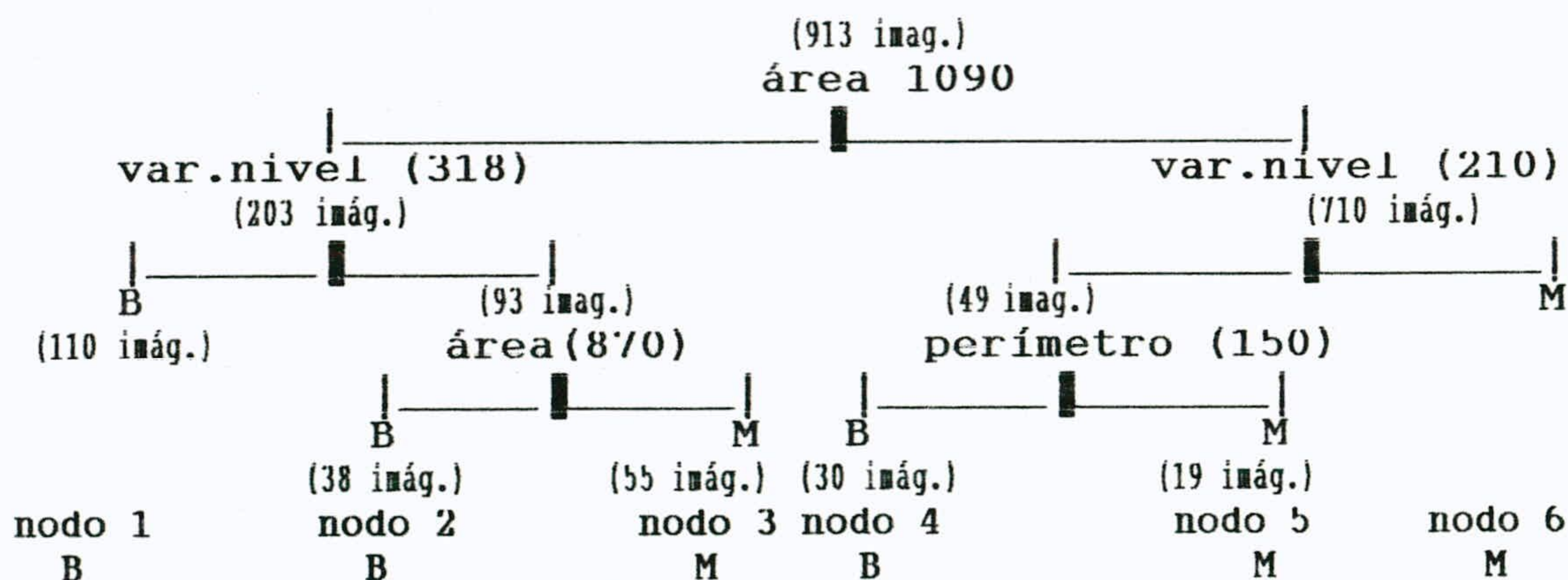


---

clasificación son idénticas en uno y otro grupo.

### III-1-1-B Resultados por imágenes con CART (22 variables)

En éste caso el árbol más simple con el menor error estimado (0.50 +/- 0.046) presentaba 6 nodos terminales. Por tanto el árbol está más ramificado que por casos y su error es casi el doble, lo que nos lleva a la conclusión que el tratamiento por imágenes es peor que por casos. Las matrices de clasificación pone de manifiesto que el error del experimento aumenta a causa de peores clasificaciones de imágenes de núcleos benignos.



Tras la selección de las imágenes adecuadas contamos con 211 imágenes con núcleos benignos y 702 imágenes con núcleos malignos.





---

MATRIZ DE CLASIFICACIÓN DE LA MUESTRA DE APRENDIZAJE

Clase Verdadera		
	B	M
B	153	25
M	58	677

MATRIZ DE PROBABILIDAD DE LA MUESTRA DE APRENDIZAJE

Clase Verdadera		
	B	M
B	0,73	0,04
M	0,27	0,96

MATRIZ DE CLASIFICACIÓN DE LA VALIDACIÓN CRUZADA

Clase Verdadera		
	B	M
B	149	44
M	62	658

MATRIZ DE PROBABILIDAD DE LA VALIDACIÓN CRUZADA

Clase Verdadera		
	B	M
B	0,71	0,06
M	0,29	0,94

Asignándole una puntuación de 100 a la más relevante, iría en éste orden:

área	100
esfericidad	98
perímetro	95
feret x	86
feret y	78
desviación estericidad	55
„ área	52
varianza niveles gris	49
circularidad	35
des. var. nivel gris	21
simetría	17
media desv. nivel gris	13
desv. perímetro	13
desv. momento 4	11
desv. simetría	11
desv. elipticidad	9
desv. feret y	8
desv. media nivel gris	8
momento 4	8
elipticidad	7
desv. feret x	6
desv. circularidad	5



---

En este caso dejan de ser más importantes las desviaciones de las variables para situarse en los primeros lugares las medias de las mismas, debido al tratamiento conjunto de las imágenes. Los valores extremos de las imágenes van compensándose por lo que perdemos el valor discriminante de las desviaciones, que sería más acentuado en los casos malignos.

También podemos observar que la varianza de los niveles de gris tiene una escasa importancia discriminadora en el orden de importancia de las variables, sin embargo en el árbol de decisión es variable decisoria en dos nodos, debido al poder clasificador del nodo anterior (área de los nucleos).

### III-1-1-C Resultados por casos con Análisis Discriminante

El grupo de variables que selecciona el programa como el que mejor predice el grupo a que cada caso pertenece es el formado por: feret de las x de la célula, varianza de niveles de gris, feret de las y y desviación del feret de las x. Este grupo define la función de clasificación:



FUNCIÓN DE CLASIFICACIÓN

	Benignas (B)	Malignas (M)
VARIABLES		
teret de las x	11.5214	12.24885
teret de las y	29.24427	30.89411
estericidad	-5.33094	-5.60955
varianza niveles	0.05815	0.07584
desv. teret x	6.52402	6.90467
CONSTANTE	-429.695	-491.3447

Definiéndose un hiperplano mediante la fórmula:

$$\text{BENIGNAS} = 11.5214 (\text{teret } x) + 29.24427 (\text{teret } y) - 5.33094 (\text{estericidad}) + 0.05815 (\text{var. niv.}) + 6.52402 (\text{desv.teret } x) - 429.695$$

$$\text{MALIGNAS} = 12.24885 (\text{teret } x) + 30.89411 (\text{teret } y) - 5.60955 (\text{estericidad}) + 0.07584 (\text{var.niv.}) + 6.90467 (\text{des.teret } x) - 491.3447$$

La Matriz de Clasificación sería la siguiente:

MATRIZ DE CLASIFICACIÓN

Grupo	% correcto	nº casos clasificados	
		B	M
B	89.4	42	5
M	93.7	6	89
Total	91.5	48	94

TEST DE CLASIFICACIÓN JACKKNIFED

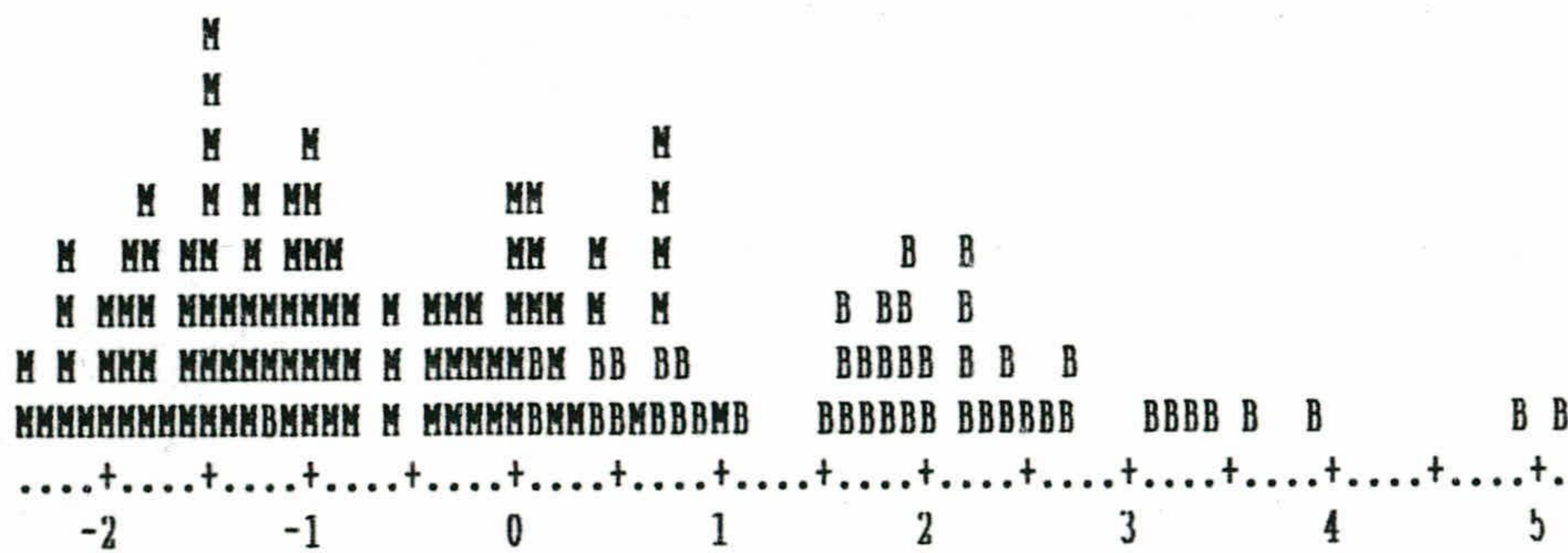
Grupo	% correcto	nº casos clasificados	
		B	M
B	85.1	40	7
M	91.6	8	87
Total	88.4	48	94



Como vemos los resultados no son muy diferentes de los obtenidos con CART pero la función de clasificación es más difícil de entender que un árbol de clasificación.

HISTOGRAMA DE LA CLASIFICACIÓN POR CASOS CON ANÁLISIS DISCRIMINANTE UTILIZANDO V. PLANIMÉTRICAS

En ella se aprecia la zona de confusión de los casos benignos y malignos aunque uno y otro grupo quedan bastante delimitados.



III-1-2 EXPERIMENTO 2: VARIABLES TEXTURALES

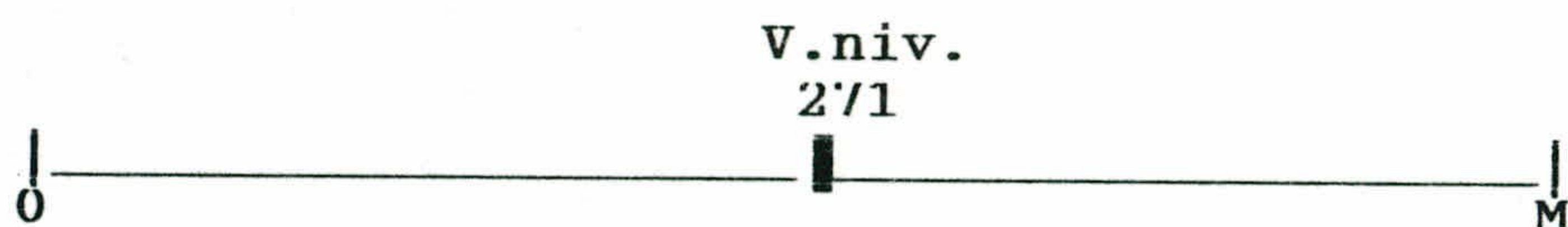
III-1-2-A Resultados por casos con CART

El árbol más simple con el menor error estimado (0.53 +/- 0.097) presentaba dos nodos terminales, siendo la mejor variable decisoria la varianza de los niveles de gris y el mejor corte en 2/1. Valores menores son considerados



---

benignos (B) y valores mayores malignos (M).



Los resultados obtenidos, aproximadamente con un error el doble más alto que los obtenidos con el primer grupo de variables tratados por casos (error  $0.26 \pm 0.071$ ), nos demuestran que éstas variables, aunque muy orientativas para el citólogo, en éste tipo de imágenes de archivo teñidas con Papanicolaou, no son excesivamente útiles para la discriminación diagnóstica.

MATRIZ DE CLASIFICACIÓN DE LA MUESTRA DE APRENDIZAJE

	Clase Verdadera	
	B	M
B	29	4
M	18	91

MATRIZ DE PROBABILIDAD DE LA MUESTRA DE APRENDIZAJE

	Clase Verdadera	
	B	M
B	0,62	0,04
M	0,38	0,96

MATRIZ DE CLASIFICACIÓN DE LA VALIDACIÓN CRUZADA

	Clase Verdadera	
	B	M
B	27	5
M	20	90

MATRIZ DE PROBABILIDAD DE LA VALIDACIÓN CRUZADA

	Clase Verdadera	
	B	M
B	0,57	0,05
M	0,43	0,95



Como podemos ver en la matriz de probabilidad del test de validación cruzada sólo el 57% de los casos (0,57) los clasifica correctamente, con los casos malignos se obtiene un alto porcentaje de bien clasificados. La aplicación de éste clasificador va a dar muchos errores de falsos positivos.

La importancia de las variables en orden decreciente, asignándole un valor de 100 a la más relevante, iría en éste orden:

varianza niveles gris	100
desviación varianza niveles de gris	76
,, contraste (coocurrencias)	57
,, media niveles gris	53
simetría	44
contraste (coocurrencias)	38
desviación máx. nivel gris (cooc.)	35
,, energía (coocurrencias)	35
,, entropía nivel gris	34
,, " " " " (dif.)	34
,, máximo " "	34
media niveles gris	29
desviación energía de los niveles	29
momento 4	27
entropía niveles (diferencias)	23
,, " "	22
diferencia-inversa (coocurrencias)	21
desviación momento 4	18
máximo (coocurrencias)	15
desviación máximo (diferencias)	15
,, entropía (coocurrencias)	15
máximo (diferencias)	15
energía (diferencias)	14
entropía (coocurrencias)	10
desviación simetría	10
máximo de los niveles	8
desviación dif.-inv. (coocur.)	7
,, energía (diferencias)	7
energía niveles	5



---

grueso de las variables con respecto a las dos primeras, es muy bajo.

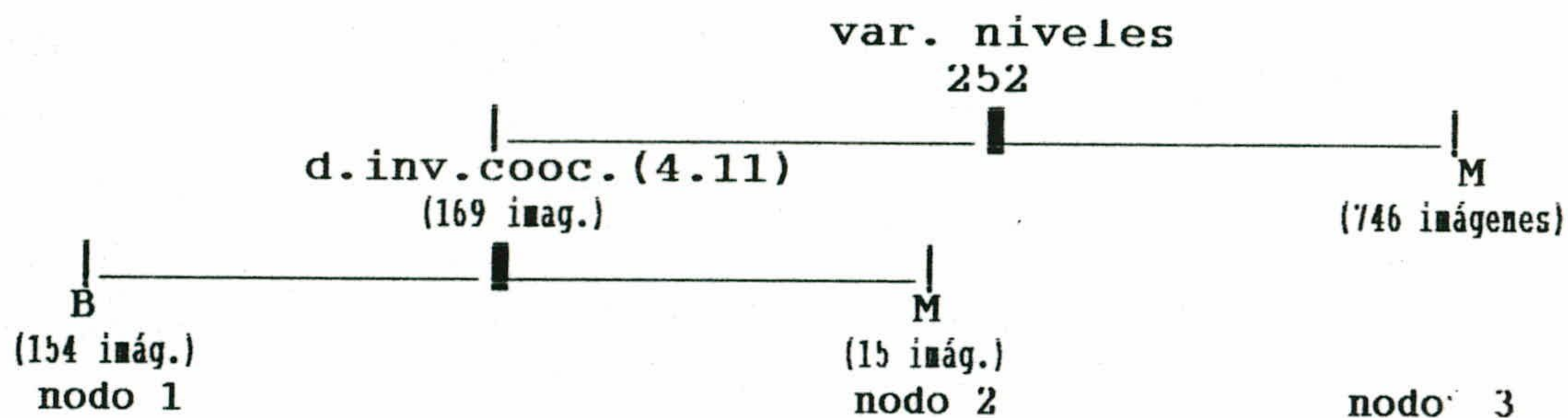
Ya que no es objetivo de éste trabajo utilizar otro método de tinción diferente al Papanicolaou, será interesante para un proyecto futuro un estudio con extensiones tratadas con tinción específica para los núcleos, con el objeto de determinar con más exactitud el poder discriminador de éstas variables texturales.

### III-1-2-B Resultados por imágenes con CART

El árbol óptimo tenía tres nodos terminales y un error relativo de  $0.79 \pm 0.055$ , lo que resulta de aproximadamente la mitad más alto que el árbol obtenido para las variables planimétricas tratadas por imágenes y de tres veces más que las mismas tratadas por casos.

En el primer nodo decisorio CART estimó como mejor variable la **varianza de los niveles de gris** con el mejor corte en 252 y en el segundo nodo decisorio la **desviación de la inversa de las coocurrencias** con el mejor corte en 4.11.





En éste caso son 211 imágenes de núcleos benignos y 702 imágenes de núcleos malignos.

MATRIZ DE CLASIFICACIÓN DE LA MUESTRA DE APRENDIZAJE

		Clase Verdadera	
		B	M
B		107	47
M		104	655

MATRIZ DE PROBABILIDAD DE LA MUESTRA DE APRENDIZAJE

		Clase Verdadera	
		B	M
B		0,51	0,07
M		0,49	0,93

MATRIZ DE CLASIFICACIÓN DE LA VALIDACIÓN CRUZADA

		Clase Verdadera	
		B	M
B		91	47
M		120	655

MATRIZ DE PROBABILIDAD DE LA VALIDACIÓN CRUZADA

		Clase Verdadera	
		B	M
B		0,43	0,07
M		0,57	0,93

En éste caso por imágenes los resultados son aún más bajos que por casos, clasificándose correctamente sólo un 43% de las benignas.



---

La importancia de las variables en orden decreciente, asignándole una puntuación de 100 a la más relevante, iría en éste orden:

varianza niveles gris	100
contraste (coocurrencias)	30
desviación dif-inv ,,	30
desviación varianza nivel gris	29
simetría	28
desviación máximo (coocurrencias)	27
media niveles de gris	26
desviación de la simetría	26
,, máximo (diferencias)	26
entropía niveles gris	24
desviación energía (coocurrencias)	23
diferencia-inversa (coocurrencias)	22
desviación entropía (coocurrencias)	22
,, ,, (diferencias)	21
,, media nivel gris	20
momento 4	19
desviación momento 4	18
entropía (diferencias)	17
energía (diferencias)	16
máximo (diferencias)	15
desviación contraste (coocurrencias)	15
,, energía niveles gris	13
energía niveles	12
entropía (coocurrencias)	11
desviación entropía niveles gris	10
,, máximo niveles gris	10
energía (coocurrencias)	9
máximo (coocurrencias)	9
máximo niveles gris	

Como vemos la varianza de niveles gris es la única que lleva el peso de la discriminación, su desviación, que tenía también importancia en las variables texturales tratadas por casos, ha perdido valor al tratarlas por imágenes por el mismo motivo que explicamos anteriormente: los valores extremos van compensándose.



---

### III-1-2-C Resultados por casos con Análisis Discriminante

---

#### FUNCIÓN DE CLASIFICACIÓN

	Benignas (B)	Malignas (M)
VARIABLES		
Var. nivel gris	0.08125	0.10995
Des. media ,,	-0.09400	0.08465
Des. simetría	75.88245	64.13486
Med. Energía	-4.98101	-3.69599
Desv. energía	0.80444	-1.27695
CONSTANTE	-89.32743	-95.10765

Definiéndose un hiperplano mediante la fórmula:

$$\text{BENIGNAS} = 0.08125 - 0.09400 + 75.88245 - 4.98101 + 0.80444 - 89.3274$$

$$\text{MALIGNAS} = 0.10995 + 0.08465 + 64.13486 - 3.69599 - 1.27695 - 95.10765$$

La Matriz de Clasificación sería la siguiente:

#### MATRIZ DE CLASIFICACIÓN

Grupo	% correcto	nº casos clasificados	
		B	M
B	80.9	38	9
M	87.4	12	83
Total	84.1	50	92

#### TEST DE CLASIFICACIÓN DE JACKKNIFED

Grupo	% correcto	nº casos clasificados	
		B	M
B	78.7	37	10
M	83.1	14	81
Total	80.9	51	91

La clasificación con Análisis Discriminante obtiene bastante mejores resultados en la correcta clasificación de los casos benignos que con CART (57%) aunque inferior en los







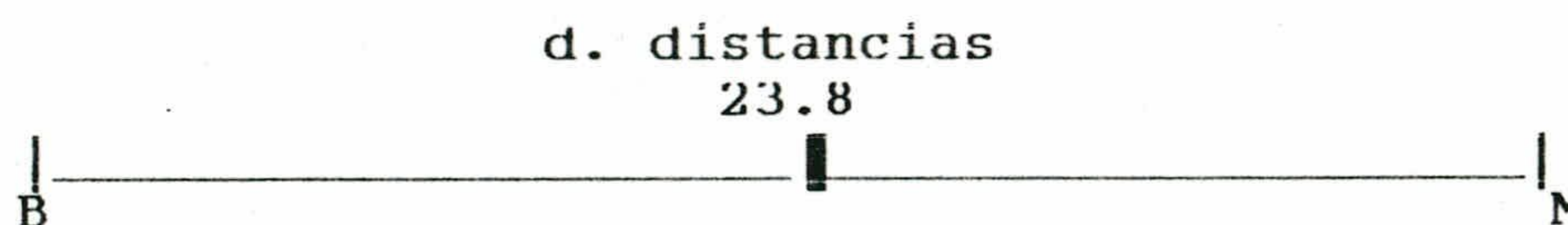
---

## III-2 RESULTADOS DEL ANALISIS DE LOS GRUPOS

### III-2-1 ANALISIS DE LAS DISTANCIAS NUCLEARES CON GRAFOS PLANARES (medidas de desorden)

#### III-2-1-A Resultados por casos con CART

La variable decisoria es la desviación de las distancias, presentando el mejor corte en 8.92 con un error relativo de  $0.09 \pm 0.042$ . Mucho más bajo que el obtenido con los mejores resultados del primer experimento (variables planimétricas).



#### MATRIZ DE CLASIFICACIÓN DE LA MUESTRA DE APRENDIZAJE

	Clase Verdadera	
	B	M
B	45	1
M	2	94

#### MATRIZ DE PROBABILIDAD DE LA MUESTRA DE APRENDIZAJE

	Clase Verdadera	
	B	M
B	0,96	0,01
M	0,04	0,99

#### MATRIZ DE CLASIFICACIÓN DE LA VALIDACIÓN CRUZADA

	Clase Verdadera	
	B	M
B	45	2
M	2	93



---

MATRIZ DE PROBABILIDAD DE LA VALIDACIÓN CRUZADA

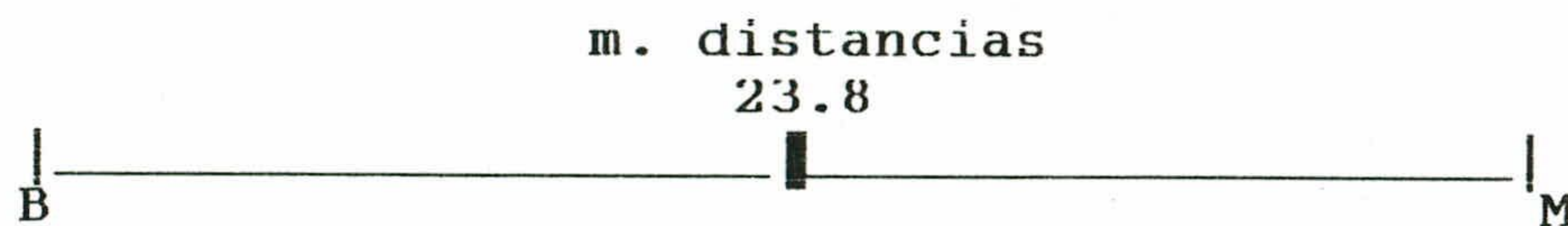
	Clase Verdadera	
	B	M
B	0,96	0,04
M	0,02	0,98

El orden de importancia de las variables sería el siguiente:

desv. de la media de las distancias	100
media de las distancias	96
máximo	74
mínimo	67

III-2-1-B Resultados con CART en los grupos individuales

El mejor árbol presenta 2 nodos terminales con un error relativo de  $0.39 \pm 0.04\%$ , la variable decisoria es la media de las distancias de los grafos con el mejor corte en 23.8, por debajo de éste valor es considerado benigno y por encima maligno. El tipo de árbol es igual al obtenido por casos pero con un error bastante más alto.



En éste caso, tras la selección adecuada de la imágenes de los grupos contamos con 150 grupos benignos y 289 grupos malignos.



---

MATRIZ DE CLASIFICACIÓN DE LA MUESTRA DE APRENDIZAJE

		Clase Verdadera	
		B	M
B	123		31
M	27		258

MATRIZ DE PROBABILIDAD DE LA MUESTRA DE APRENDIZAJE

		Clase Verdadera	
		B	M
B	0,82		0,11
M	0,18		0,89

MATRIZ DE CLASIFICACIÓN DE LA VALIDACIÓN CRUZADA

		Clase Verdadera	
		B	M
B	123		31
M	27		258

MATRIZ DE PROBABILIDAD DE LA VALIDACIÓN CRUZADA

		Clase Verdadera	
		B	M
B	0,82		0,11
M	0,18		0,89

Como podemos observar también en éste experimento, la desviación de la variable nos discrimina mejor por casos frente a la media que nos discrimina mejor cuando el experimento se realiza con los grupos individuales, y siempre con errores superiores.

El orden de importancia de las variables en orden decreciente sería:

media de las distancias de los nucleos	100
máximo de las distancias	86
mínimo	75
desv. media	66



---

### III-2-1-C Resultados por casos con Análisis Discriminante

#### FUNCIÓN DE CLASIFICACIÓN

VARIABLES	Malignas (M)	Benignas (B)
Media distancias	0.83779	0.52996
Desv. media , ,	1.97573	0.40200
CONSTANTE	-28.09963	-6.57970

Definiéndose un hiperplano mediante la fórmula:

$$\text{MALIGNAS} = 0.83779(\text{Med. dis.}) + 1.97573(\text{desv. med.}) - 28.09963$$
$$\text{BENIGNAS} = 0.52996(\text{Med. dis.}) + 0.40200(\text{desv. media}) - 6.57970$$

La Matriz de Clasificación sería la siguiente:

#### MATRIZ DE CLASIFICACIÓN

Grupo	% correcto	nº casos clasificados	
		B	M
B	95.7	45	2
M	98.9	1	94
Total	97.3	46	96

#### TEST DE CLASIFICACIÓN DE JACKKNIFED

Grupo	% correcto	nº casos clasificados	
		B	M
B	95.7	45	2
M	98.9	1	94
Total	97.3	46	96

Los resultados son similares a los obtenidos con CART en el análisis por casos.

Aplicando las fórmulas de la Eficacia (Et),



Aplicando las fórmulas de la Eficacia (Et), Sensibilidad (S), Especificidad (Es), Valor predictivo (VPP) positivo y valor predictivo negativo (VPN) de los datos obtenidos en la Matriz de Clasificación de la Validación Cruzada, obtenemos para éste experimento:

$$Et = (vp + vn) / total * 100 = 97.9$$

$$S = (vp / (vp + fn)) * 100 = 98.9$$

$$Es = (vn / (vn + fp)) * 100 = 95.7$$

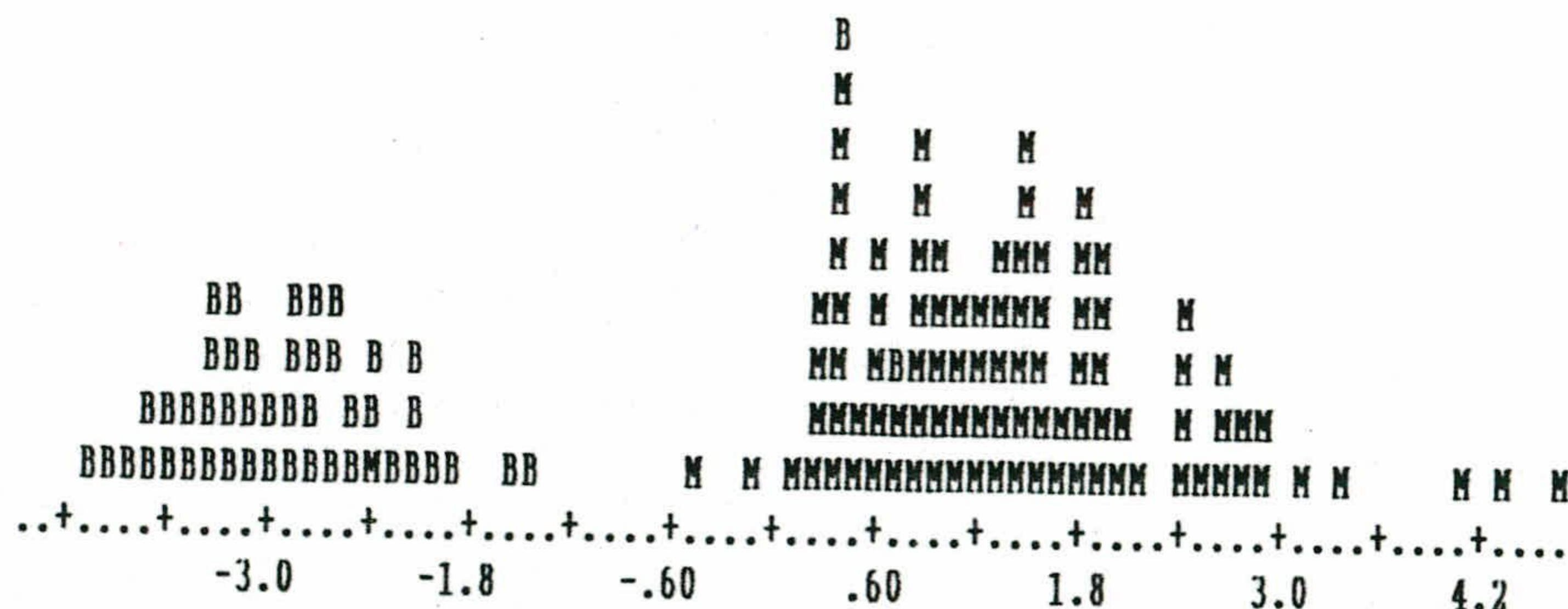
$$VPP = (vp / (vp + fp)) * 100 = 97.9$$

$$VPN = (vn / (vn + fn)) * 100 = 97.8$$

Siendo:

vp = verdaderos positivos = 94  
 vn = verdaderos negativos = 45  
 fp = falsos positivos = 2  
 fn = falsos negativos = 1

HISTOGRAMA DE LA CLASIFICACIÓN POR CASOS CON ANÁLISIS DISCRIMINANTE UTILIZANDO GRAFOS PLANARES



Como vemos las clases están casi completamente definidas lo que se traduce en un bajo error de clasificación.



III-2-1-D Resultados con Análisis Discriminante en los grupos individuales

FUNCION DE CLASIFICACION

VARIABLES	Benignas (B)	Malignas (M)
Media distancias	0.05413	0.18550
Mínimo	0.52995	0.62740
CONSTANTE	-5.11950	-10.47030

Definiéndose un hiperplano mediante la fórmula:

$$\begin{aligned} \text{Benignas} &= 0.05413 (\text{Med. dis.}) + 0.52995 (\text{mín. dis.}) - 5.11950 \\ \text{Malignas} &= 0.18550 (\text{Med. dis.}) + 0.62740 (\text{mín. dist.}) - 10.47030 \end{aligned}$$

La Matriz de Clasificación sería la siguiente:

MATRIZ DE CLASIFICACIÓN

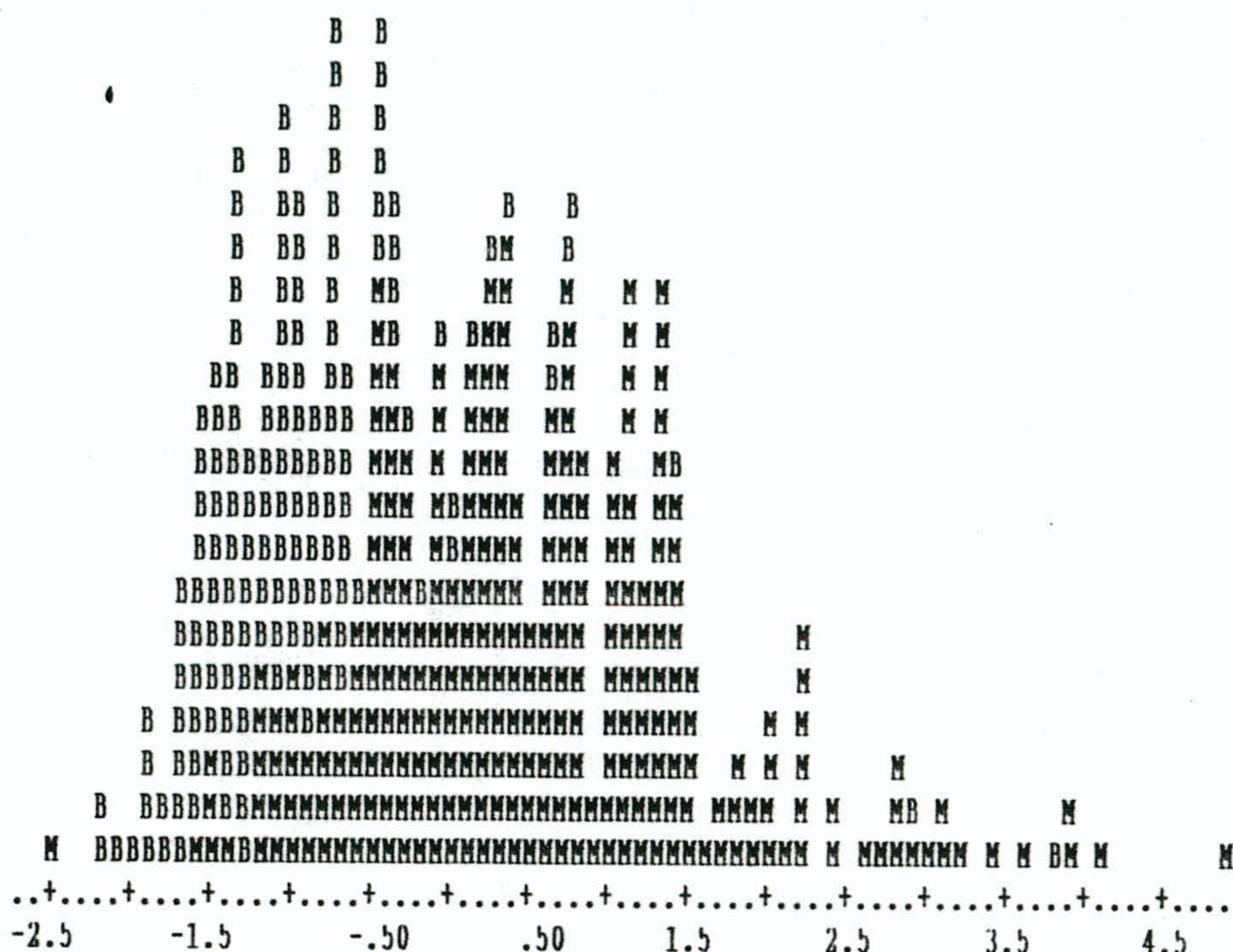
Grupo	% correcto	nº casos clasificados	
		B	M
B	90.7	136	14
M	73.7	76	213
Total	82.2	212	227

TEST DE CLASIFICACIÓN DE JACKKNIFED

Grupo	% correcto	nº casos clasificados	
		B	M
B	90.7	136	14
M	73.7	76	213
Total	82.2	212	227



HISTOGRAMA DE LA CLASIFICACIÓN POR GRUPOS INDIVIDUALES CON ANÁLISIS DISCRIMINANTE UTILIZANDO GRAFOS PLANARES



Si comparamos el porcentajes de correctas clasificaciones (90.7% de B y 73.7% de M) con el realizado agrupandolos por casos y Análisis Discriminante (95.7% de B y 98.9% de M), podemos observar que los primeros son muy inferiores, disminuyendo el poder discriminador a costa de los casos malignos. Este hecho también puede verse gráficamente en ambos Histogramas de Clasificación: en el análisis por casos las clases benignas y malignas están claramente separadas, lo que no ocurre en el análisis por grupos individuales. La interpretación se da en el sentido de que también las distancias nucleares en los grupos de un mismo caso de carcinoma son muy heterogéneas frente a la



---

mayor homogeneidad de las distancias nucleares en los grupos de un mismo caso de lesión benigna y la razón es la misma que la dada para el análisis de los nucleos: el análisis global de los grupos individuales compensa los valores extremos por lo que disminuye el poder discriminador de las variables.

Los resultados nos indican que no sólo existe heterogeneidad de los nucleos dentro de los grupos de un caso de carcinoma mamario sino que también lo es el desorden de los nucleos entre los grupos, todo ello frente a una mayor homogeneidad de los casos benignos.



---

**DISCUSION**



---

## IV DISCUSION

Para conseguir el objetivo de este trabajo: construcción de un clasificador de extensiones citológicas de lesiones mamarias utilizando análisis de imagen, había que resolver primeramente tres problemas: el primero era la SELECCIÓN DE LOS RASGOS CITOLÓGICOS CUANTIFICABLES, el segundo EL TRATAMIENTO DE LOS MISMOS PARA LA METODOLOGÍA ELEGIDA y el tercero EL ANÁLISIS de ellos.

### IV-1 CONSTRUCCION DEL CLASIFICADOR

#### IV-1-1 SELECCION DE LOS RASGOS CITOLOGICOS

Hay numerosos trabajos que ilustran la variabilidad en el diagnóstico anatomopatológico por interpretación subjetiva de los rasgos, ya sea por un mismo observador o por varios (Baak 1982 b, Delides 1982, Holman 1983, Stenkvist 1983, Ooms 1983). Progresivamente se va imponiendo la necesidad de su cuantificación para hacerlos más objetivos fundamentalmente en aquellos casos en que las características no están completamente definidas, lo que denomina Galera-Davidson (1991) cuadros que no tienen un "patrón concreto".

La mayoría de los trabajos de cuantificación morfométrica en lesiones mamarias se centran en encontrar parámetros relevantes que orienten el pronóstico (Walgren 1976,



---

Zajdela 1979, Van Bogaert 1980, Baak 1982 y 1985, Maehle 1982 y 1988, Linden 1985) o el grado nuclear o histológico (Cornelisse 1985, Diest 1989, Salmon 1991), en su mayoría parámetros nucleares estudiados en cortes histológicos. Los trabajos morfométricos con el fin de discriminar lesiones benignas de lesiones malignas también se centran fundamentalmente en el núcleo, ya sea sobre cortes histológicos (Sandritter 1974, Bhattacharjee 1985) o sobre extensiones citológicas (Gadenne 1985, Marcos 1986, Wittekind 1987, King 1988, Detweiler 1988, Hutchinson 1991, Wolberg 1993).

Este trabajo ha tratado de abarcar el mayor número de variables que el citólogo valora para llegar a un diagnóstico de malignidad o benignidad. Para ello se basó en un protocolo que se confeccionó (Protocolo I) y que contenía todos aquellos rasgos que el citólogo valora en la misma secuencia que él suele seguir: desde los rasgos más generales, que son analizados con un objetivo de bajo aumento, hasta los detalles más específicos que son analizados a mayor aumento. Escuetamente SON: RASGOS GENERALES DEL EXTENDIDO, RASGOS DE LOS GRUPOS, RASGOS ESPECÍFICOS DE LOS NUCLEOS.

a - Rasgos generales del extendido:

Las tumoraciones malignas tienen un crecimiento rápido y anárquico que conduce a la destrucción de los tejidos produciendo necrosis tumoral e inflamación. Además



---

las células malignas tienen la particularidad de tener menos cohesión entre ellas y por tanto en los extendidos hay más células desgajadas de los grupos que en los extendidos benignos (Jimenez-Ayala 1988). Estos rasgos no eran fácilmente cuantificables con técnicas de Análisis de Imagen y , además, existen factores de confusión que obligan continuamente al experto a seleccionar, lo que va en contra del objetivo de éste trabajo donde se persigue la mayor automatización y objetividad posible. Así la necrosis tiene unas características especiales de color y textura que el experto discrimina de fondos muy sucios (inflamados o fibrinosos) o muy teñidos pero que en imágenes en niveles de grises pueden ser idénticos. Por otro lado valorar las células sueltas semejantes a la de los grupos, obligaría a una continua selección y a un contaje arduo pues puede haber muchos otros tipos de células sueltas como hematies, leucocitos , células del estroma...

**b - Rasgos de los grupos:**

En el cáncer, al existir menos cohesión celular, los grupos de células de los ductus en los extendidos son más pequeños y el crecimiento anárquico se traduce en desorden celular, solapamiento de los núcleos dentro del grupo y alteración de la relación área nuclear/área citoplasmática. Estos rasgos eran más cuantificables y objetivos. El desorden de los núcleos integrados y el solapamiento de ellos se



---

abordaron automáticamente con técnicas de Análisis de Imagen en los grupos pequeños. Para la relación entre el área nuclear y el área citoplasmática de las células, se realizó una inferencia en los grupos con técnicas de análisis semiautomáticas. Y para la información que pudiera derivarse del tamaño y forma de los grupos, se escogieron los dos grupos más grandes que había en todos los extendidos de un mismo caso.

c - Rasgos específicos de los núcleos:

Son los rasgos más objetivos ya que los núcleos tienen unos contornos bien definidos lo que permitió que la segmentación y análisis se automatizara totalmente. Los rasgos de los nucleolos no se tuvieron en cuenta de forma explícita, pues la tinción de Papanicolaou no los tiñe selectivamente y muchos acúmulos cromatínicos pueden ser confundidos como tales, sus características quedaron inmersas dentro de las variables texturales. Los rasgos nucleares se dividieron en dos grupos: VARIABLES PLANIMÉTRICAS O GEOMÉTRICAS (área, perímetro, ejes, factor de forma, factor elipsoidal...etc) y VARIABLES TEXTURALES, resultando un total de 50 variables.



El segundo problema era tratar imágenes tridimensionales como bidimensionales. Algunos trabajos, cuyo material son cortes histológicos, realizan inferencias volumétricas (Nielsen 1988, Artacho 1993) siguiendo la metodología descrita por Gundersen y Jensen (1985), pero esto no es posible con las extensiones citológicas, técnica rápida, económica, fácil e inocua, al contrario que el sistema para obtener cortes histológicos.

Las imágenes obtenidas a partir de extensiones no es exactamente la correspondiente a un corte a ese nivel: existe un emborronamiento que corresponde a porciones desenfocadas de los planos inferior y superior, el denominado "ruido fuera de foco". Muchos autores han abordado éste problema de diversas maneras, algunos basándose en las propiedades físicas del microscopio como en la convolución del objeto con una función de emborronamiento (psf) y en la función de transferencia óptica del microscopio en el dominio de Fourier (OTF) (Erhardt 1985, Streibl 1984 y 1985, Diaspro 1990), y otros minimizando el ruido usando filtros para la reconstrucción de la imagen 3-D (tres dimensiones) a partir de la 2-D obtenida en los distintos planos (Macias-Garza 1988 y 1989). Sin embargo ésta metodología tiene un alto coste computacional y de almacenamiento, además de ser complicada y laboriosa, lo que la inutilizaría para el objetivo de éste



---

trabajo.

Siendo que la metodología seguida por el experto en citología es la de observar la mayoría de los rasgos en el plano de enfoque más adecuado, aceptamos la aproximación de tratar las imágenes como 2-D.

#### IV-1-3 METODOLOGIA PARA EL ANALISIS DE IMAGENES

Por último el tercer problema consistía en encontrar el método que automatizara lo más posible el análisis de las imágenes. Para ello se contó con la colaboración del Grupo de Análisis de Imágenes del Departamento Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada con objeto de desarrollar un software que tratara automáticamente, con técnicas de Análisis de Imagen, todas las variables que el experto seleccionó como importantes y que eran susceptibles de ser cuantificables. El proceso y los resultados era controlado por nosotros para que se ajustaran a los criterios de diagnóstico citológico. Así se desarrolló un **Sistema Informático de Tratamiento Celular** que dejó resueltos muchos de los problemas que plantea el análisis de imagen: la homogeneización de las imágenes para hacerlas independientes de las condiciones de captación, el realce y tratamiento de éstas para que, finalmente, los algoritmos de automatización pudiesen actuar correctamente.



---

## IV-2 EL CLASIFICADOR

Algunos trabajos que aparecen en la literatura se han centrado en la consecución de un clasificador para la patología mamaria: King en 1988 utiliza un programa de análisis textural (el Leitz TAS-plus) y Análisis Discriminante para el estudio de tres variables nucleares que selecciona por su superior discriminación: la SD del factor de forma, la media de la condensación cromatínica y la SD de las manchas oscuras en el núcleo. Detweiler (1988) y Hutchinson (1991) desarrollan un clasificador utilizando variables nucleares, tanto planimétricas como texturales (éstas limitadas al histograma de niveles de gris) y variables planimétricas de los grupos (área de éstos, distancia entre grupos, polaridad de ellos) utilizando el analizador de imagen Micro Vax-based y analizando los datos con Análisis Discriminante. Wolberg (1987 y 1989) desarrolla un clasificador con árboles binarios usando el FACT (Fast and Automatic Classification Tree) utilizando rasgos cualitativos de todo el extendido a los que aplica una escala arbitraria del 1 al 10 y posteriormente en 1993 (Wolberg 1993) un clasificador que utiliza variables nucleares estudiadas con análisis de imagen utilizando el método lineal de análisis de datos MSM (Multisurface Method of Pattern Separation).

Los programas utilizados en éste trabajo realizan la clasificación de las lesiones en base a rasgos de los



---

nucleos (geométricos y texturales) y de los grupos (desorden de los nucleos y áreas). Sin embargo la variable relación área nuclear/área citoplasmática de las células, importante en la valoración de la malignidad, no pudo ser automatizada a causa de la dificultad en realizar la segmentación y a la imposibilidad de hacerla sobre células independientes, por lo que se procedió a estimarla sobre el grupo en su conjunto con tratamiento semiautomático. Para ello se desarrolló una MACRO con el programa VISILOG (1988) que trata y analiza el área de interés interactivamente (Herrera 1993, Marcos 1994).

En una exposición secuencial comentaremos los tres puntos en los que éste trabajo se ha esforzado en hacer alguna aportación en comparación con respecto los trabajos existentes: en primer lugar el material recogido para LA BASE DE DATOS, en segundo lugar EL MÉTODO DE SEGMENTACIÓN AUTOMÁTICO Y SEMIAUTOMÁTICO, en tercer lugar LAS VARIABLES TRATADAS DE NUCLEOS Y GRUPOS Y en último lugar EL ANÁLISIS DE LOS DATOS CON ÁRBOLES BINARINARIOS (METODOLOGÍA CART) VERSUS ANÁLISIS DISCRIMINANTE.

#### IV-2-1 LA BASE DE DATOS

La gran cantidad de material del presente trabajo: 95 casos malignos y 47 benignos (142 casos de lesiones mamarias) recogidos en alrededor de 1750 imágenes de nucleos y grupos y otras tantas de sus fondos lo que hacía un total aproximado de 3500 imágenes con 3550 nucleos y 850 grupos,



---

debía dar variabilidad a la base de datos ya que recogía diversos tipos de cáncer de mama (de células pequeñas y de células grandes, más diferenciados o menos, más o menos pleomórficos...etc), y si bien existe pocas clases diferentes al Carcinoma Ductal (8%) es porque también son poco frecuentes (el periodo de recogida de casos es de 5 años).

#### IV-2-2 LA SEGMENTACION AUTOMATICA Y SEMIAUTOMATICA

La segmentación manual es difícil y laboriosa y por ello sólo aplicable en trabajos de investigación. Agilizar el proceso de segmentación es requisito indispensable para pretender que éste tipo de técnica pueda ser de ayuda en la rutina diagnóstica. Los algoritmos de segmentación automáticos conseguidos a través del proyecto de investigación conjunto al que nos hemos referido representan ese paso necesario desde la metodología de investigación a la metodología diagnóstica.

Los algoritmos de segmentación de los núcleos implementados en éste trabajo mejoran la segmentación manual, no sólo en rapidez, lo que es obvio, sino en la precisión para delimitar las regiones de interés en las que efectuar las medidas. El trazado manual es desigual y subjetivo (Marchevsky 1987), no depende sólo de la discriminación visual de cada uno de los posibles usuarios sino también de su pulso. De ésta forma dos personas no harían nunca igual un



---

pulso. De ésta forma dos personas no harían nunca igual un trazado sobre el borde de un núcleo e, incluso, es poco probable que una misma persona pueda repetirlo igual. Sobre esta cuestión podemos citar el trabajo de Barry y cols. (1985) que analiza la concordancia de 3 parejas de observadores realizando medidas morfométricas (área, perímetro y factor de forma) sobre núcleos en cortes histológicos de cáncer de mama utilizando una tableta digitalizadora. La reproducibilidad fué buena para el área pero progresivamente peor para las dos siguientes variables.

Sobre el algoritmo de grafos planares hay escasos -- antecedentes de utilización de algo similar para el diagnóstico biomédico. La valoración del desorden en los núcleos lo realiza el experto en base a su experiencia. En los trabajos sobre extendidos citológicos que consideran ésta variable, la estimación se realiza de forma cualitativa (Dziura 1979, Wolberg 1987 y 1989), y su papel en la discriminación de las lesiones es muy relevante. Tener una herramienta que la cuantifique podía ser muy útil, y nos remitimos a nuestros resultados en los que obtenemos en el análisis por casos una Eficiencia del 97,9%, Sensibilidad del 94.7 %, Especificidad del 85.1 %, Valor Predictivo Positivo (VPP) del 98.9 % y Valor Predictivo Negativo (VPN) del 95.7 % , cifras que son similares a las obtenidas por diversos autores con la citología por punción aspiración de lesiones mamarias (Garrido-Teruel 1978, Jimenez-Ayala 1988, Fernández-Izquierdo 1989)



---

que, como ya se ha mencionado, es uno de los métodos diagnósticos más fiables.

La macro realizada para obtener la relación área nuclear/área citoplasmática (AN/AC) de los grupos se basa en técnicas de Morfología Matemática y la segmentación se obtiene por binarización interactiva. Ambas herramientas empiezan a ser habituales en análisis de imagen. Existen escasos trabajos que investiguen ésta variable en extensiones de lesiones mamarias (Boon 1982, Wolber 1987 y 1989, García-Valdecasas 1988).

#### V-2-3 LAS VARIABLES TRATADAS

Las variables fueron tratadas por imágenes y por casos. El análisis por casos daba mejores resultados ya que ponía de manifiesto la heterogeneidad de las poblaciones celulares, rasgo importante en las lesiones malignas, resultando la desviación de las variables más importantes como discriminadores de las lesiones que sus medias. El análisis por imágenes compensaba los valores extremos por lo que era la media de las variables las que llevaba el peso de la discriminación.

Los trabajos cuantimétricos de lesiones mamarias que describe la literatura obtienen la información fundamentalmente del análisis de los núcleos, siendo la media



---

del área nuclear el más habitualmente recogido en las publicaciones (Boon 1982, Marcos 1986, Wittekind 1987, Detweiller 1988, Hutchinson 1991, Wolberg 1993) seguido por el perímetro nuclear (Boon 1982 y Wittekind 1987, Hutchinson 1991). Son escasos los trabajos que recogen análisis textural (Sandritter 1974, Detweiller 1988, Hutchinson 1991, Wolber 1993), morfometría de los grupos (Detweiller 1988, Hutchinson 1991) o morfometría del citoplasma de la célula (Boon 1982, García-Valdecasas 1988). Realizándose la selección del área de interés de forma manual o semiautomática.

Existe ya mucho software de Análisis de Imágenes de aplicación en biomedicina (remitirse al apartado 1-4-4) y para la patología mamaria con fines diagnósticos van apareciendo algunos trabajos de clasificación de las lesiones utilizando técnicas de segmentación semiautomáticas (Gadenne 1985, King 1988, Detweiler 1988, Hutchinson 1991, Wolberg 1993), siendo prácticamente inexistentes los trabajos con técnicas de segmentación automáticas.

En cortes histológicos de 41 casos de lesiones mamarias Gadenne (1985) construye un clasificador con la selección de 5 variables entre un total de 27. Estas son: media del área, perímetro,  $\text{perímetro}^2/\text{área}$ , error estándar del total de las granulaciones y el  $\text{perímetro}^2/\text{área}$  de éstas. Para ello utilizó la tinción estequiométrica de Feulgen, un mínimo de 100 núcleos por caso y el sistema de análisis de



---

imágen Leitz TAS-plus. Testado el clasificador con los mismos casos, se clasificaba correctamente el 90% de los malignos y el 71% de los benignos. Al compararlos con nuestros resultados con CART encontramos que en ambos casos el área y el perímetro tienen alto poder discriminador. El resto de las variables no se han tenido en cuenta ya que no se ha hecho análisis granulométrico ni densitométrico.

Detweiller (1988) construye un clasificador basado en parámetros nucleares geométricos (área, perímetro y factores de forma) y texturales (medidas basadas en el histograma de niveles de grises), y un clasificador basado en parámetros de los grupos (número de éstos por campo, distancia entre ellos, área de los grupos y media de niveles de grises de éstos). Para ello utiliza 5 casos benignos y 5 casos malignos de extensiones citológicas teñidas con hematoxilina-eosina para la muestra de aprendizaje (189 núcleos en total entre benignos y malignos y 81 campos, cada campo incluía de 1 a 5 grupos celulares) y añadió 8 casos de difícil diagnóstico citológico para testar el clasificador. El contorno nuclear lo dibuja manualmente con una tableta gráfica. Resultando discriminate en los nucleos individuales: la media del área y las variables texturales, con una correcta clasificación en el 88,9% de los casos. En el presente trabajo (con CART), también la media del área es el estimador decisorio cuando los datos se tratan por imágenes y la desviación del área cuando se tratan por casos, con una correcta clasificación



---

con Análisis Discriminante del 92.3 para las variables geométricas y del 85 % para las variables texturales. De las variables de los grupos estudiados al azar encuentran en los casos malignos: grupos más pequeños, menos celulares y más numerosos con una clasificación correcta del 77.8% en los benignos y el 88.9% de los malignos y cuando se unían ambos clasificadores el 100%. En el éste trabajo el clasificador con los grupos de células se basa en medidas de desorden, clasificando correctamente el 95.7% de los benignos y el 98.9% de los malignos (valores sensiblemente superiores). No se ha considerado las variables: número de grupos por campo y distancia entre ellos porque, aunque en las lesiones malignas los grupos de células suelen ser más pequeños y numerosos, es un factor poco objetivo ya que depende mucho de la cantidad de material extraído. La tercera variable, área de los grupos, se ha considerado de dos formas: para relacionarla con el área nuclear de los mismos y para analizar el tamaño de los grupos grandes, no encontrando diferencias en ninguna de las dos estimaciones realizadas, probablemente debido a que tanto los grupos grandes como los grupos pequeños no fueron escogidos al azar. En cualquier caso la relación de áreas en los grupos sería interesante estudiarla con respecto al número de núcleos que contiene. Ambas metodologías son interactivas. Queda como nuevo objetivo de investigación aunar todos los clasificadores en uno solo, esperando que también mejoren los resultados.



---

Posteriormente Hutchinson (1991) realiza un estudio similar en 18 casos de lesiones benignas y 43 malignas, pero utilizando la tinción estequiométrica de Feulgen. Encuentra el perímetro nuclear primera variable en importancia discriminadora, seguida de las demás que encuentra Detweiler (1988) y del factor de forma. Los resultados que obtiene con ambos clasificadores son inferiores, así como la unión de ambos. Las diferencias puede deberse a las diferentes tinciones en uno y otro trabajo.

King (1988) realiza también un estudio de discriminación entre grupos de lesiones mamarias en 39 extensiones citológicas teñidas con de Feulgen y segmentando previa binarización interactiva. El estudio lo centra en parámetros nucleares geométricos (área, perímetro, factor de forma), densitométricos y granulométricos, obteniendo para cada uno de ellos: la media, la desviación estándar y el rango intercuartil. Obteniendo con Análisis Discriminante, para el grupo benignas versus malignas, una clasificación correcta del 100% de los casos benignos y del 78.6 % de los malignos, con el grupo de variables formado por: la media del perímetro nuclear, el intercuartil de la marginación y la desviación del índice de DNA. En nuestro caso la segmentación de los nucleos es completamente automática (aunque su metodología es similar a la utilizada por nosotros para la segmentación de los grupos) y clasifica correctamente el 89.4 de las benignas y el 93.7 de las malignas con las variables:



---

feret x, feret y, esfericidad del núcleo, desv. de los niveles de gris y desv. feret x. El trabajo de King y el nuestro presentan pocas variables comunes (área, perímetro, factor de forma y sus SD) pero ninguna de ellas coinciden en ambos clasificadores estadísticos. La explicación puede deberse a las diferentes técnicas de tinción utilizadas y a los diferentes tamaños de muestra utilizados.

El clasificador con árboles binarios de Wolberg (1987 y 1989) de extendidos citológicos teñidos con hematoxilina-eosina utiliza las variables cualitativas: densidad celular en los grupos, cohesión marginal de éstos, tamaño del grueso de las células epiteliales (comparadas con los eritrocitos), cantidad de núcleos aislados, cromasia nuclear (comparada con leucocitos y linfocitos), aspecto nucleolar, presencia de mitosis, relación núcleo-citoplasmática celular, número de núcleos desnudos, la uniformidad en el tamaño y forma de las células epiteliales y la presencia de células apocrinas, inflamatorias, mioepiteliales. De ellas selecciona como rasgos decisivos la densidad y solapamiento de los núcleos en el grupo y en segundo lugar la cohesión de las células en la zona marginal de los grupos, obteniendo una Sensibilidad del 98%, Especificidad del 94% y VPP del 92% . Todas estas variables coinciden mayoritariamente con las recogidas en éste trabajo, resultando también como clasificador óptimo el basado en el desorden nuclear de los grupos (concretado en la media de la desviación de las distancias) y con resultados



---

algo superiores (sensibilidad 98.9%, Especificidad 95.7%, VPP 97.8% ), además de ser variables cuantificables y por tanto mejor reproducibles y menos subjetivas. Posteriormente, en 1993, Wolberg publica un nuevo trabajo de clasificación en 119 lesiones mamarias (de las que elige aleatoriamente 569 imágenes), aunque ya utilizando análisis de imagen con técnica de segmentación interactiva. En éste caso elige 11 variables nucleares: tamaño nuclear, área, varianza radial, compactividad ( $\text{perímetro}^2/\text{área}^3$ ), concavidad, varianza de los niveles de gris, varianza del tamaño, el mayor tamaño y la mayor compactividad, consiguiendo una clasificación correcta en el 90% de las imágenes, utilizando un clasificador lineal. Nosotros obtenemos para las variables nucleares planimétricas pero analizadas por casos, un 92.3% de bien clasificados (ya concluimos que el análisis por casos daba mejores resultados que por imágenes).

Boon (1982) en 77 casos de lesiones mamarias, realiza mediciones (nucleares y citoplasmáticas) segmentando interactivamente sobre células aisladas o sobre las que se encuentran en los bordes de los grupos en extensiones citológicas, no encontrando diferencias significativas para la media de la relación área nuclear/área citoplasmática pero sí para la desviación estándar. Sí aprecian diferencias García-Valdecasas y Herlinger (1988), los cuales estudiaron el área nuclear y citoplasmática de 100 células de cada grupo de lesiones mamarias en extensiones citológicas (10 casos de



---

Fibroadenomas, 10 casos de MFQ, 10 casos de Papiloma y 34 carcinomas de mama), aunque no explican el tipo de tinción utilizada ni el método de selección de las células. Las medidas las realizan con un microproyector y el planímetro de Haaf, encontrando que la media de ambas áreas aumenta en las lesiones con ésta secuencia: Fibroadenoma, Mastopatía Fibroquística, Papiloma, siendo el carcinoma el que presenta una mayor media del área nuclear (no consiguen obtener el área media citoplasmática en las lesiones malignas por su dificultad). Nosotros no estudiamos ésta variable sobre las células individuales, lo que daría un estimador menos sesgado pero, en contraposición, las células aisladas se alteran más fácilmente por la desecación o la tinción que los grupos de células. Segundo, no todos los extendidos tienen suficientes células aisladas intactas por lo que habría de recurrirse, como hace Boon en su trabajo, a las células de los bordes de los grupos. Esto implica que si bien por el borde externo queda bien delimitado el contorno de la célula, por los otros bordes debe realizarse un corte subjetivo donde se estime esté el citoplasma de la célula elegida y separarla de las demás segmentando manualmente. Wolberg en 1987 y 1989 no encuentra ésta variable decisoria al construir un árbol de clasificación con diversos rasgos valorados cualitativamente. En nuestro caso el clasificador que utiliza la relación de áreas a través de los índices I e I' (remitirse al apartado II-4-1) no discrimina las lesiones. Queda como nueva línea de investigación relacionar los índices de cada grupo con el



---

número de núcleos integrantes, procedimiento más interactivo que escapa al objetivo de este trabajo.

Resumiendo las diferencias: el clasificador óptimo de éste trabajo se centra en parámetros de desorden, es completamente automático, es decir muy rápido, se ha utilizado un gran número de muestra de aprendizaje, lo que se traduce en una gran versatilidad para el diagnóstico de nuevos casos, y con todo ello presenta una gran Eficiencia, Sensibilidad y Especificidad.

El resto de los trabajos morfométricos cuyo objetivo es la discriminación de lesiones benignas-malignas mamarias que recoge la bibliografía se centran en parámetros nucleares. Así Wittekind (1987) encuentra discriminante los tres parámetros estudiados: media del área nuclear, media del diámetro máximo y la media del perímetro realizando análisis univariante en las mediciones obtenidas de extensiones citológicas. Además hace un estudio de comparación entre morfometría realizada con ocular micrométrico y analizador de imagen, encontrando una baja correlación en los resultados (67%). Sobre corte histológico también Bhattacharjee (1985) obtiene que el área y el eje mayor son discriminantes utilizando ocular micrométrico. En nuestro clasificador con variables planimétricas, también estos dos parámetros ocupan los primeros lugares en orden de importancia.



---

Puesto que los diferentes rasgos texturales son muy tenidos en cuenta por el experto, es propósito para nuevas líneas de investigación trabajar con extensiones recientes teñidas con Papanicolaou y con tinciones estequiométricas específicas del núcleo. Las tinciones específicas tienen poca o ninguna afinidad por otros elementos diferentes al específico y por ello pueden utilizarse para estimar la concentración de la sustancia (sobre éste punto nos remitimos al apartado I-5-1-A que habla de la Citofotometría). Sin embargo, los frotis teñidos con la tinción de Papanicolaou están influidos por multitud de factores que modifican el resultado, entre ellos la presencia de mayor o menor cantidad de otras sustancias con afinidad por el mismo colorante. Además, las tinciones históricas, como las utilizadas en nuestro caso, pierden paulatinamente definición y color.

#### IV-2-4 METODOLOGIA PARA EL ANALISIS DE LOS DATOS

Si comparamos ambos métodos de clasificación, CART y Análisis Discriminante, no hay una clara ventaja de uno sobre el otro en los resultados pero, sin embargo, los árboles binarios son más fáciles de usar e interpretar que las expresiones algebraicas del Análisis Discriminante. A las mismas conclusiones llegan otros trabajos de investigación en los que se han comparado ambas metodologías (Aus 1986, Fernández-Valdivia 1991).



---

La mayoría de los trabajos de investigación analizan sus resultados con métodos estadísticos, entre ellos se incluyen la mayoría de los trabajos morfométricos en lesiones mamarias a los que se ha hecho referencia anteriormente. Pero más frecuentemente los investigadores están haciendo uso de otro tipo de clasificadores como son los árboles de clasificación (clasificador de aprendizaje). Así en lesiones mamarias tenemos los clasificadores con árboles de decisión de Wolberg de 1987 y 1989 o de gradación tumoral realizada por Wolberg en 1988 o el trabajo con Redes Neuronales y Análisis Discriminante para la gradación tumoral de Dawson en 1991.



---

## V LINEAS FUTURAS

1- Para una profundización en el análisis de las características necesarias en los procesos de clasificación de imágenes citológicas, se hace necesario la búsqueda de procedimientos para el tratamiento de imágenes 3-D lo que con lleva problemas de estudio de algoritmos para la búsqueda de rasgos en este tipo de imágenes tridimensionales.

En la investigación de rasgos citológicos importantes discriminadores de lesiones malignas queda pendiente para nuevos proyectos de investigación con otros objetivos:

2- Los rasgos texturales, utilizando extensiones tratadas con tinciones específicas estequiométricas.

3- La relación Area nuclear/area citoplasmática de los grupos en relación al número de núcleos visibles por grupos.

4- Aunar el poder discriminante de los Clasificadores de variables planimétricas nucleares y Clasificador con gratos planares con el fin de determinar la capacidad clasificatoria global.



---

## VI CONCLUSIONES

1 Tras la observación meticulosa y pormenorizada del material citológico ha sido posible priorizar los rasgos citológicos que permiten clasificar a las lesiones mamarias como benignas o malignas.

2 La transformación inicial de las imágenes digitalizadas mediante la ecualización del histograma (al restar a la imagen original su fondo) permite una homogeneización en el tratamiento de las citologías de forma que se ayuda a la automatización de los procesos de segmentación, necesarios en la detección de características.

3 El análisis de los resultados de las distintas variables agrupadas en casos es más discriminador que el análisis por imágenes individuales (resultando la desviación de las variables más discriminadoras por casos y la media por imágenes individuales).

4 La clasificación en base a distancias entre los núcleos (medidas de desorden) realizada con los grafos planares resultan ser muy discriminantes y con altos valores de Sensibilidad, Especificidad y Eficiencia (en concreto: la desviación de la media de las distancias en primer lugar, cuando se analizan los grupos por casos y la media de las distancias cuando se analizan por grupos individuales).



---

5 En la clasificación en base a variables nucleares, las variables planimétricas dan también buenos resultados llevando el peso de la discriminación el área, la estericidad y los ejes (ya sean sus medias o sus SD según sea el tratamiento por imágenes individuales o agrupadas por casos). Las variables texturales en trotis históricos teñidos con la tinción de Papanicolaou tienen bajo poder discriminador, centrándose éste en la varianza de niveles de grises o su SD.

6 La clasificación en base a las áreas del grupo no tienen ningún poder discriminador.

7 Comparando los resultados con CART y con Análisis Discriminante hay escasas diferencias, pero el análisis con CART resulta de mejor comprensión y de más fácil interpretación.

8 Se ha conseguido automatizar diversas variables útiles en la discriminación de las lesiones malignas de mama (en su defecto se ha aplicado análisis semiautomático a las imágenes).

9 Los algoritmos de segmentación automática aplicados funcionan rápida y eficazmente como lo demuestran los resultados obtenidos con las Matrices de Clasificación.



---

## BIBLIOGRAFIA

- 1 Adams LR, Dahlgren SE. Cytophotometric measurements of the DNA content of lung tumors. *Acta Pathologica Microbiologica Scandinavica*. 1968; 72: 561-574
- 2 Alty JL, Coombs MJ. *Sistemas Expertos. Conceptos y ejemplos*. Madrid. Ed. Diaz de Santos. 1986.
- 3 Anderson DE. A genetic study of human breast cancer. *Journal of the National Cancer Institute*. 1972; 48: 1029-1034.
- 4 Anderson DE, Badzioch D. Combined effect of family history and reproductive factors on breast cancer risk. *Cancer*. 1989; 63: 349-353.
- 5 Artacho E, Roldán R, Chicano MT, López A, Ruiz FJ. Análisis estereológico nuclear en lesiones endometriales. *Actas del VIII Congreso Nacional de Histología*. Jaén: Editado MA Peinado y JA Pedrosa. 1993: 139-140.
- 6 Atlas del cancer en España. Vitoria: Ediciones G López, A Escobar, M Errezola, 1984: 91-97.
- 7 Aus HM, Harms H, Haucke M, Beritova J, Meulen ter V. Statistical evaluation of computer markers to detect leukemias. *Pattern Recognition Letters*. 1986; 4: 231-241.
- 8 Azúa, J. *Citología por punción con aguja fina*. Barcelona. Ed. Salvat. 1987: 194-215.
- 9 Azzopardi JG. Problems in Breast Pathology. En: Bennington JI. *Major Problems in Pathology*. London-Philadelphia-Toronto. WB Saunders Company. 1979; 11.
- 10 Baak JPA, Kurver PHJ, de Graaf S, Boon ME. Morphometry for prognosis prediction in breast cancer. *The Lancet*. 1981; 8: 315.
- 11 Baak JPA, Dop van H, Kurver PHJ, Hermans J. The value of morphometry to classic prognosticators in breast cancer. *Cancer*. 1985; 54: 374-382.
- 12 Baak JPA, Lindeman J, Overdiep SH, Langley FA. Disagreement of histopathological diagnosis of different pathologists in ovarian tumours. *European Journal of Obstetrics, Gynecology and Reproductive Biology*. 1982 b; 13: 51-55.
- 13 Baak JPA, Kurver PHJ, de Snoo-Niewlaat AJE, Makkink B, Boon ME. Prognostic indicator in breast cancer - morphometric methods. *Histopathology*. 1982; 6: 327-339.



- 
- 14 Balaban-Malenbaum G, Grove G, Gilbert F. Increased DNA contents of HSR-marker chromosomes of human neuroblastoma cells. *Experimental Cell Research*. 1979; 119: 419.
- 15 Bentura ML, Colmenar L. Fundamentos y principios de los métodos planimétricos y citofotométricos. *Oncología*. 1988; 11: 30-38.
- 16 Berryman IL, Harvey JM, Sterrett GM, Papadimitriou JM. The nuclear DNA content of human breast carcinoma. Associations with clinical stage, axillary lymph node status, estrogen receptor status and outcome. *Analytical and Quantitative Cytology and Histology* 1987; 9: 429-434.
- 17 Bhattacharjee DK, Harris M, Faragher EB. Nuclear morphometry of epitheliosis and intraduct carcinoma of the breast. *Histopathology*. 1985; 9: 511-516.
- 18 Bishop JM. Oncogenes. En: *El Cáncer*. Primera edición. Barcelona: Graefesa, Prensa Científica, 1985: 62-74.
- 19 Bloom HJG, Richardson WW. Histologic grading and prognosis in breast cancer *British Journal of Cancer*. 1957: 359-377.
- 20 BMDP. Estadistical Software Manual. PMDP SOSFT inc. Los Angeles (California). 1990.
- 21 Boon ME, Trott PA, van Kaam H, Kurver PJH, Leach A, Baak JPA. Morphometry and Cytodiagnosis of breast lesions. *Virchows Archiv*. 1982; 396: 9-18.
- 22 Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. 1ª ed. Los Angeles (California). Wadsworth Inc. 1984.
- 23 Brugal G. Action concertée européenne: cytologie automatique et analytique. *Innov. Tech. Biol. Med*. 1986; 7: 115-123.
- 24 Brugarolas A, Aparicio LM. El desarrollo terapeutico: Cauce de investigación y progreso de la oncología médica. VI Reunión Científica del Fondo de Investigaciones Sanitarias de la Seguridad Social (FISss). Madrid: Instituto Nacional de la Salud. 1989: 45-65.
- 25 Cady M D. Familial bilateral cancer of the breast. *Annals of Surgery*. 1970; 172: 264.
- 26 Cairns J. El problema del cáncer. En: *El Cáncer*. 1ª edición. Ed. Prensa Científica. Barcelona. 1985: 12-23.
- 27 Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*
-



---

(PAMI). 1986; 6: 679-699.

28 Carlile T. Breast cancer detection. *Cancer*. 1981; 47: 1164-1169.

29 Carroll KK, Hopkins GJ. Dietary polyunsaturated fat versus saturated fat in relation to mammary carcinogenesis. *Lipids*. 1979; 14: 155.

30 Carroll KK, Khor HT. Effects of level and type of dietary fat on mammary tumors induced in female Sprague-Dawley rats by 7,12-dimethylbenz(a)anthracene. *Lipids*. 1971; 6: 415.

31 CART (TM). Versión 1.0. California Statistical Software Inc. California (USA). 1984.

32 Charney J, Moore DH. Neutralization of murine mammary tumor virus by sera of women with breast cancer. *Nature*. 1971; 229: 627.

33 Clark GM, Sledge GW Jr, Osborne CK, McGuire WL. Survival from first recurrence: relative importance of prognostic factors in 1015 breast cancer patients. *Journal of Clinical Oncology*. 1987; 5: 55-61.

34 Connors RW, Harlow CA. Toward a structural textural analyzer based on statistical methods. In *Image Modeling*. 1<sup>a</sup> ed. New York. A. Rosenfeld eds. Academic Press. 1980: 29-39.

35 Consensus Meeting. Is "Fibrocystic Disease" of the breast Precancerous? *Archives of Pathology and Laboratory Medicine*. 1986; 110: 171-173.

36 Cornelisse CJ, van de Velde CJ, Caspers RJ, Moolenaar AJ, Hermans J. DNA ploidy and survival in breast cancer patients. *Cytometry*. 1987; 8: 225-234.

37 Cornelisse CJ, van Driel-Kulker AMJ, Meyer F, Ploem S. Automated recognition of atypical nuclei in breast cancer cytology specimens by iterative image transformations. *Journal of Microscopy*. 1985; 137: 101-110.

38 Coulter W. High-speed automatic blood cell counter and cell size analyzer and reader. *Proc Natl Electronics Conf (Chicago)*. 1956: 1034-1042.

39 Chalkley HW., Cornfield J, Park H. A method for estimating volume-surface ratios. *Science*. 1949; 110: 295.

40 Chicharro Molero JA, Burgos Rodriguez R. Test Diagnósticos. En: *Metodología de Investigación y Escritura Científica*. Granada. Escuela Andaluza de Salud Pública. 1994: 143-152.



- 
- 41 Christov K, Milev A, Todorov V. DNA aneuploidy and cell proliferation in breast tumors. *Cancer*. 1989; 64: 673-679.
- 41 Crocker J, Nar P. Nucleolar organizer regions in Lymphomas. *Journal of Pathology*. 1987; 151: 111-118.
- 42 Davis LS, Mitiche A. Edge detection in textures. In *Image Modeling*. 1<sup>a</sup> ed. Londres. A. Rosenfeld eds. Academic Press. Inc. 1980: 95-102.
- 43 Davis BW, Gelber RD, Goldhirsch A, et al. Prognostic significance of tumor grade in clinical trials of adjuvant therapy for breast cancer with axillary lymph node metastasis. *Cancer*. 1986; 58: 2662-2670.
- 44 Dawson AE, Austin RE Jr, Weinberg DS. Nuclear grading of breast carcinoma by image analysis. Classification by multivariate and neural network analysis. *American Journal of Clinical Pathology*. 1991; 95: 29-37.
- 45 de Torres IM. Aplicación de técnicas de análisis de imagen en la valoración del grado de malignidad de los tumores vesicales. Tesis doctoral. Publicacions de la Universitat Autònoma de Barcelona 1988.
- 46 de Torres I, Moragas, García M, Sans M, Allende H. Estudio textural del segmento conservado de pared coronaria en el infarto de miocardio por ateromatosis coronaria. *Patología*. 1986; 19:94-97.
- 47 de Ward F, Cornelis JP, Aoki K, et al. Breast cancer incidence according to weight and height in two cities of the Netherlands and in Aichi Prefecture, Japan. *Cancer*. 1977; 40: 1269.
- 48 de Ward F. The bimodal age distribution of patients with mamary carcinoma. *Cancer*. 1964; 17: 141-151.
- 49 Delesse A. Procède mecanique pour determiner la composition des roches. *Comptes Rendus de l'Academie des Sciences (Paris)*. 1984; 25: 544-548.
- 50 Delides GS, Garas G, Georguli, et al. Intralaboratory variations in the grading of breast carcinoma. *Archives of Pathology and Laboratory Medicine*. 1982; 106: 126-128.
- 51 Detweiler R, Zahniser DJ, García GL, Hutchison M. Contextual analysis complements single-cell analysis in the diagnosis of breast cancer in fine needle aspirates. *Analytical and Quantitative Cytology and Histology*. 1988; 10: 10-15.
- 52 Diamond LW, Nathwani BN, Rappaport H. Flow cytometry in the diagnosis and classification of malignant lymphoma leukemia. *Cancer*. 1982; 50: 1122-1135.
-



- 
- 53 Diaspro A, Adami M, Nicoli C. Three dimensional representation of biostructures imaged with an optical microscope. *Image and Vision Comput.* 1990; 8: 130-139.
- 54 Diest van PJ, Risse EKJ, Schipper NW, Baak JPA. Comparison of microscopic grading and morphometric features in cytological breast cancer specimens. *Pathology, Research and Practice.* 1989; 185: 612-616.
- 55 Dmochowski L, Seman G, Gallasher HS. Viruses as possible etiologic factors in human breast cancer. *Cancer.* 1969; 24: 1241.
- 56 Dantorth DN. Enfermedades de la mama. En: Dantorth DN. *Tratado de Obstetricia y Ginecología.* 4ª edición. Madrid. Editorial Interamericana. 1986.
- 57 Doussal V, Tubiana-Hulin M, Fridman S, Hacene K, Spyrtos F, Brunet M. Prognostic value of histologic grade nuclear components of Scarff-Bloom-Richardson. *Cancer.* 1989; 64: 1914-1921.
- 58 Doust BD, Doust VL. Imágenes ultrasónicas, radiológicas y con marcadores radiactivos en el diagnóstico obstétrico. En: Dantorth DN. *Tratado de Obstetricia y Ginecología.* 4ª Ed. Madrid. Ed. Interamericana. 1986: 541-543
- 59 Dowle CS, Owainati A, Robins A, Burns k, Ellis IO, Elston CW, Blamey RW. Prognostic significance of the DNA content of human breast cancer. *British Journal of Surgery.* 1987; 74: 133-136.
- 60 Dressler L, Seamer LC, Owens M A, Clark G M, McGuire WL. DNA Flow Cytometry and Prognostic factors in 1331 frozen breast cancer specimens. *Cancer* 1988; 61: 420-427.
- 61 Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. *The New England Journal of Medicine.* 1985; 312: 146.
- 62 Dziura BR, Bontiglio TA. Needle Cytology of the breast. A quantitative and qualitative study of the cells of benign and malignant ductal neoplasia. *Acta Cytologica.* 1979; 23: 332-340.
- 63 Egan LR. Mamografía, Xerorradiografía y Termografía. *Clinicas Obstétricas Ginecológicas.* Junio 1976.
- 64 Enchev VG, Dobрева PV. Cytophotometric measurements of nuclear DNA content values in human breast lesions. *Arch Geschwulstforsch.* 1987; 57: 31-37.
-



- 
- 65 Erhardt A, Zinser G, Komitowski D, Billie J. Reconstructing 3-D light-microscopic images by digital image processing. *Applied optics*. 1985; 24: 194-200.
- 66 Esquivias JJ, Herrera C, Mendoza J, Ruiz-Cabello F. Valoración del contenido de DNA en fibroblastos humanos procedentes de cultivo. Comparación entre citometría de flujo y estática, usando técnicas de Feulgen y Hematoxilina progresiva. *Patología*. 1994; 27: 23-29.
- 67 Fawcett DW. La glándula mamaria. En: *Tratado de Histología*. 11ª edición. Madrid. Interamericana. 1989: 909-920.
- 68 Fernandez-Cid A. *Patología mamaria*. Barcelona. Editorial Salvat. 1982.
- 69 Fernandez-Cid A, Sentis J, Porteiro E, Petracco A, Gonzalez R. Edad y cáncer de mama. *Acta Ginecológica*. 1979; 35: 315-318.
- 70 Fernández-Valdivia J. Algunas técnicas para el procesamiento de imágenes microscópicas. En: *Algunos aspectos del tratamiento de la información en Inteligencia Artificial*. Granada. Servicio de Publicaciones de la Universidad de Granada. 1991: 24 -261.
- 71 Fernández-Valdivia, J. *Identificación Automática de Rasgos en Imágenes Microscópicas*. Tesis Doctoral. Universidad de Granada. 1991.
- 72 Ferrer Roca O. Olga<sup>2</sup>scan software de análisis textural con aplicaciones citológicas. II. Receptores estrogénicos. *Citología*. 1989; 11: 129-136.
- 73 Ferrer-Roca O. *Análisis de imagen (I): Principios básicos de tratamiento y evaluación estadística*. Santa Cruz de Tenerife. Editorial Confederación de Cajas de Ahorros. 1986.
- 74 Ferrer-Roca O. *Análisis de Imagen (II): aplicaciones*. Santa Cruz de Tenerife. Editorial Confederación de Cajas de Ahorros. 1990.
- 75 Ferti-Passantonopoulou AD, Panani AD. Common cytogenetic finding in primary breast cancer. *Cancer Genetics and Cytogenetics*. 1987; 27: 289.
- 76 Fogelman F, Mejia C. *Third International Symposium on Knowledge Engineering*. Madrid. 1988.
- 77 Gadenne Cl, King EB, Chew KL, Duarte LA, Mayall BH, Petrakis NL. A quantitative study of duct cells from benign and malignant breast disease. *Acta Cytologica*. 1985; 29: 917-818.
-



---

78 Galera Davidson H. Introducción sobre las necesidades de los métodos cuantimétricos y sus aplicaciones. Actas del XV Congreso Nacional de Anatomía Patológica. Sevilla. 1991.

79 Galera Davidson H. La citometría interactiva en patología. Actas del XV Congreso Nacional de Anatomía Patológica. Sevilla. 1991.

80 Gallagher HS. Early phases in the development of breast cancer. Cancer. 1969; 24: 1170.

81 Galloway M. Texture analysis using gray level run lengths. Computer Graphics Image Processing. 1974; 4: 172-199.

82 Gammal EB, Carroll KK, Plunkett ER. Effects of dietary fat on mammary carcinogenesis by 7,12-dimethylbenz(a)anthracene in rats. Cancer Research. 1967; 27: 1737.

83 García-Valdecasas R, Herlinger RM. Papel de la citometría en el frotis mamario. Toko-Ginecología Práctica. 1988; 47: 77-81.

84 Garrido A, Pérez de la Blanca N. Segmentación de imágenes citológicas. Technical Report CCIA. 1994; 94: 3-6.

84b Garrido A, Pérez de la Blanca N, Fernández-Valdivia J. Segmenting non-random noise biomedical digital images. Proc. of the VI Simposium Nacional de Reconocimiento de formas y análisis de imágenes. 1995: 434-442.

85 Garrido-Teruel R, Marques Bravo A, Novo Domínguez A, Gonzalez Navarro G. Correlación histopatológica de los diagnósticos clínicos, mamográficos y citológicos sobre la malignidad de las lesiones de mama. Toko-Ginecología Práctica. 1978; 417: 221-228.

86 Giménez Mas JA, Sanz Moncasi P, Remon Garijo L. Aplicaciones de la morfometría a la citopatología: Análisis computerizado de la estructura nuclear por medio de técnicas de Morfología Matemática. Citología. 1991; 12: 101-107.

87 Greig EDW, Gray ACH. Note on the lymphatic glands in sleeping sickness. Br Med J. 1904; 1: 1252-1256.

88 Guerin M, Barrois M, Gabillot M, Riou G. Expression des proto-oncogenes c-myc et c-erbB-2 dans les cancers du sein: Implications cliniques. Bulletin du Cancer. 1989; 76: 175-180.

89 Gundersen HJG, Jensen EB. Stereological estimation of the volume-weighted mean volume of arbitrary particles observed on random sections. Journal of Microscopy. 1985; 138: 127-142.



- 
- 90 Haagensen LD, Bolian C, Haagensen DE. Breast carcinoma risk and detection. Filadelfia. WB Saunders. 1981.
- 91 Hally AD. A counting method for measuring the volume of tissue components in microscopical sections. Quart J Micr Sci. 1964; 105: 503.
- 92 Hall PA, Crocker J, Watts A, Stansfield AG. A comparison of nucleolar organizer region staining and Ki-67 immunostaining in non-Hodgkin's lymphoma. Histopathology. 1988; 12: 373-381.
- 93 Haralick RM. Statistical and structural approaches to texture. Proceedings of the IEEE. 1979; 7: 786-804.
- 94 Henderson IG, Harris JR, Kinne DW, Hellman S. Cancer of the Breast. En: De Vita VT Jr, Hellman S, Rosenberg SA. Cancer. Philadelphia. 3ª edición. Editorial JB Lippincott Company. 1984: 1197-1199.
- 95 Herrera C, Esquivias J, Marcos C, Rodriguez L. Carcinoma de Mama: variables en relación a su incidencia (grupo sanguíneo, edad, localización y antecedentes familiares). Toko-Ginecología Práctica. 1993; 52: 87-91.
- 96 Hirohata T, Nomura AM, Hankin JH, Kolotel LN, Lee J. An epidemiologic study on the association between diet and breast cancer. Journal of the National Cancer Institute. 1978; 78: 595.
- 97 Holman CDJ, Matz LR, Finlay-Jones LR, et al. Inter-observer variation in the histopathological reporting of Hodgkin's disease: an analysis of diagnostic subcomponents using kappa statistics. Histopathology. 1983; 7: 399-407.
- 98 Hrubec Z, Boice JD, Monson RR, Rosenstein M. Breast cancer after multiple chest fluoroscopies: second follow-up of Massachusetts women with tuberculosis. Cancer Research. 1989; 49: 229-234.
- 99 Howson C, Kinne D, Wynder E. Body weight, serum cholesterol and stage of primary breast cancer. Cancer. 1986; 58: 1804-1813.
- 100 Hutchinson ML, Isenstein LM, Zahniser DJ. High-Resolution and contextual analysis for the diagnosis of fine needle aspirates of breast. Analytical and Quantitative Cytology and Histology. 1991; 13: 351-355.
- 101 Hunter T. Proteínas de Oncogenes. En: El Cáncer. 1ª edición. Ed. Prensa Científica. Barcelona. 1985: 75-85.
- 102 Ing R, Hoe JHC, Petrakis NL. Unilateral breastfeeding and breast cancer. Lancet. 1977; 2: 1124.
-



- 
- 103 Jackson D, Severance A. Cytological study of nipple secretions. *Diagnosis of breast lesions. Journal of Medicine.* 1949; 41: 512.
- 104 Jimenez Ayala M. Citología por P.A.A.F. de las lesiones malignas de la glándula mamaria. *Citología.* 1988; 10: 38-52.
- 105 Kajikawa H, Kawamoto K, Herz F, Woley RC, Hirano A, Koss LG. Flow through cytometry of meningiomas and cultured meningioma cells. *Acta Neuropathologica.* 1978; 44: 183-187.
- 106 Kallioniemi OP, Blanco G, Alavaikko M, et al. Tumour DNA ploidy as an independent prognostic in breast cancer. *British Journal of Cancer.* 1987; 56: 637-642.
- 107 Kallioniemi OP, Blanco G, Alavaikko M, et al. Improving the prognostic value of DNA Flow Cytometry in breast cancer by combining DNA index and S-phase fraction. A proposed clasification of DNA histograms in breast cancer. *Cancer.* 1988; 62: 2183-2190.
- 108 King EB, Chew KL, Duarte L, Hom JD, Mayall BH, Miller TR, Petrakis NL. Image Cytometric classification of premalignant breast disease in fine needle aspirates. *Cancer.* 1988; 62: 114-124.
- 109 King MC, Go RCP, Elston RC, et. al. Allele increasing susceptibility to human breast cancer may be linked to the glutamate-pyruvate transaminase locus. *Science.* 1980; 208: 405.
- 110 Kolata G. Oncogenes give breast cancer prognosis. *Science.* 1987; 235: 160-161.
- 111 Koss LG, Woyke S, Olszewski W. Biopsia por aspiración. Interpretación citológica y bases histológicas. Buenos Aires. Editorial Médica Panamericana S.A. 1988
- 112 Kvale G, Heuch I. A prospective study of reproductive factors breast cancer. Age at first last birth. *American Journal of Epidemiology.* 1987; 126: 842.
- 113 Landeweerd GH, Gelsema ES, Bins M, Halie Mr. Interactive pattern recognition of blood cells in malignant lymphomas. *Pattern Recognition.* 1981; 14: 239-244.
- 114 Larsimont D, Kiss R, d'Olne D, de Launoit Y, et al. Correlation between nuclear cytomorphometric parameters and estrogen receptor levels in breast cancer. *Cancer.* 1989; 63: 2162-2168.
-



- 
- 115 Larsimont D, Kiss R, d'Olne D, de Launoit Y, Mattheiem W, Paridaens R. Relationship between computerized morphonuclear image analysis and histopathologic grading of breast cancer. *Analytical and Quantitative Cytology and Histology*. 1989; 11: 433-439.
- 116 Lawry J, Giri D, Rogers K, Duncan JL. The value of assessing cell proliferation in breast cancer. *Journal of Microscopy*. 1990; 159: 265-275.
- 117 Ledley RS, Ruddle FH. Chromosome analysis by computer. *Scientific American*. 1966; 214: 40-46.
- 118 Levine MD. *Vision in Man and Machine*. 1<sup>st</sup> ed. New York. McGraw Hill. 1980.
- 119 Levitt SH, Mandel J. Breast irradiation and future risk of carcinogenesis. *Frontiers of Radiation Therapy and Oncology*. 1983;17: 131.
- 120 Lin YK, Fu KS. Automatic classification of cervical cells using a binary tree classifier. *Pattern Recognition*. 1983; 16: 69-80.
- 121 Linden van der HC, Baak JPA, Lindeman J, Hermans J. Morphometry and breast cancer II. Characterisation of breast cancer cells with high malignant potential in patients with spread to lymph nodes: preliminary results. *Journal of Clinical Pathology*. 1986; 39: 603-609.
- 122 Linden van der HC, Baak JP, Lindeman J, Hermans J, Meyer CJ. Prospective evaluation of prognostic value of morphometry in patients with primary breast cancer. *Journal Clinical Pathology*. 1987; 40: 302-306.
- 123 Lin YK. Fu KS. Automatic classification of cervical cells using a binary tree classifier. *Pattern Recognition*. 1983; 16: 69-80.
- 124 López FJ, Acevedo MI, Jaramillo MA. Redes Neuronales: Evolución Histórica. *Mundo Electrónico*. 1989; 197: 57-70.
- 125 Luijendijk H. Automatic threshold selection using histograms based in the count of 4-connected regions. *Pattern Recognition Letters*. 1991; 12: 219-228.
- 126 Lynch HT. *Genetics and Breast Cancer*. Van Nostrand Reinhold. New York. 1981.
- 127 Lynch Ht, Hirayama Takeshi. *Genetic epidemiology of cancer*. Florida (USA). Ed. CRC Press, inc. 1989: 290-332.
-



- 
- 128 Lynch HT, Conway TA, Lynch JF. Phenotypic variation in hereditary breast cancer. Cancer control implications. Archives of Surgery. 1994; 129: 806-813.
- 129 Macias-Garza F, Bovick AC, Diller KR, Aggarwal SJ. Digital reconstruction of three-dimensional serially sectioned optical images. IEEE Transactions on Pattern Recognition and Artificial Intelligence. 1988; 36: 1067-1075.
- 130 Macias-Garza F, Diller KR, Bovick AC, Aggarwal SJ. Obtaining a solid model from optical serial sections. Pattern Recognition. 1989; 22: 577-586.
- 131 Maehle BO, Thoresen S, Skjaerven R, Hartveit F. Mean nuclear area and histological grade of axillary node tumour in breast cancer, related to prognosis. British Journal of Cancer. 1982; 46: 95-100.
- 132 Maehle BO, Skjaerven R, Collett K. Prognosis in node negative breast cancer patients: The importance of tumour diameter and mean nuclear area. European Journal of Surgical Oncology. 1988; 14: 21-26.
- 133 Maillet M. Epitelios Glandulares: histología e histofisiología humanas. 1ª ed. Madrid. Editorial AC. 1980.
- 134 Marcos C. Citología por punción de los tumores epiteliales malignos de mama. Tesis Doctoral. Facultad de Medicina. Universidad Autónoma de Madrid. 1986
- 135 Marcos C, Herrera C, Esquivias J, Herruzo A. Image analysis as classification method of breast cancer and benign lesions applied to fine needle aspiration cytological smears. Actas del IV Congreso Internacional de Ginecología. Montreal (Canadá). 1994: 35.
- 136 Marchevsky AM, Gil J, Jeanty H. Computerized interactive morphometry in pathology: Current instrumentation and methods. Human Pathology. 1987; 18: 320-331.
- 137 Marr D, Hildreth E. Theory of edge detection. Proceedings of the Royal Society of London. Series B: Biological Sciences. 1980; 207: 187-217.
- 138 Martin HE, Ellis EB. Aspiration biopsy. Surgery Gynecology and Obstetrics. 1934; 59: 578-589.
- 139 Martin, H., Voss, K. Computerized classification of gliomas by automated microscope picture analysis. Acta Neuropathologica. 1982; 58: 261-268.
- 140 Takahashi M. Atlas en Color de Citología del Cáncer. Editorial Científico-Médica. Barcelona. 1973.
-



- 
- 141 Matheron, G. Elements pour une Theorie des Milieux Poreux. Masson et Cie. Editeurs. Paris. 1967
- 142 Mokhtarian F, Mackworth A. Scale based description of planar curves and two dimensional shapes. IEE transactions. 1986; PAMI 8: 34-43.
- 143 Moragas A. Cuantificación de texturas microscópicas. XV Congreso Nacional de la Sociedad Española de Anatomía Patológica. Sevilla. 1991.
- 144 Moragas A. Técnicas de Análisis Textural. XII Congreso Nacional de la Sociedad Española de Anatomía Patológica. Patología (supl. esp.). 1985: 102-109.
- 145 Moragas A, Huguet P, Toran N, Rona V. Morphogenesis of pulmonary venous occlusive disesse in a newborn. Pathology, Research and Practice. 1983; 176: 176-184.
- 146 Moragas A. Las técnicas de análisis de imagen y su aplicación en Anatomía Patológica. Medicina Clínica. 1984; 83: 373-376.
- 147 Mortalidad en Andalucía 1975-1980. Editado por el Servicio de Información y Evaluación Sanitaria. 1987.
- 148 Mosteller F, Tukey JW. Data Analysis and Regression. 2ª ed. Londres. Ed Adison Wesley. 1977.
- 149 Mui JK, Fu KS. Automated classification of nucleated blood cells using a bynary tree classifier. IEEE. Transactions on Pattern Analysis and Machine Intelligence. 1980; PAMI-2: 429-440.
- 150 Muss HB, Kute TE, Case L, et al. The relation of flow cytometry to clinical and biologic characteristics in women with node negative primary breast cancer. Cancer. 1989; 64: 1894-1900.
- 151 Nielsen K, Berild GH, Bruun E, Jorgensen PE, Weis N. Stereological estimation of mean nuclear volume in prostatic cancer, the reproducibility and the possible value of estimations on repeated biopsies in the course of disease. Journal of Microscopy. 1989; 154: 63-69.
- 152 Novak ER, Woodruffi JD. Ginecología y Obstetricia. 3ª edición. Madrid. Editorial Alhambra. 1970.
- 153 Ooms ECM, Anderson WAD, Alons CL, Boon ME, Veldhuizen EW. Analysis of the perfomance of pathologists in the grading of bladder tumors. Human Pathology. 1983; 14: 140-143.
-



- 
- 154 Owainati AAR, Robins RA, Hinton C, et al. Tumour aneuploidy, prognostic parameters and survival in primary breast cancer. *British Journal of Cancer*. 1987; 55: 449-454.
- 155 O.M.S. Ginebra. Clasificación histológica de los tumores de la mama. 1982.
- 156 Palmer JO, Mcdivitt RW, Stone KR, Rudloff MA, Gonzalez JG. Flow Cytometric analysis of breast needle aspirates. *Cancer*. 1988; 62: 2387-2391.
- 157 Paul, J., von Goldamme, E. (1990). Neural net applications in Medicine. En: Symbols versus Neurons?. J. Stender, T. Addis, Eds. IOS Pres. 1990.
- 158 Pérez de la Blanca N. Métodos de segmentación automática de imágenes biomédicas. Proyecto de cátedra. Granada. 1993
- 159 Pérez-Iraola M. Análisis Morfométrico en Citología Urinaria: Bases para la creación de un sistema experto para aplicarlo al diagnóstico precoz del cáncer. Tesis Doctoral. Universidad de Granada. 1989.
- 160 Pienta KJ, Coffey DS. Correlation of nuclear morphometry with progression of breast cancer. *Cancer*. 1991; 68: 2012-2016.
- 161 Pike MC, Henderson BE, Krailo MD, Duke A, Roy S. Breast cancer in young women and use of oral contraceptives: posible modifying effect of formulation and age and use. *Lancet*. 1983; 22.
- 162 Preparata FP, Shamos ML. Computational Geometry: An introduction. 1ª ed. Berlín. Addison-Wesley. 1987.
- 163 Preston, K.Jr.(1961). Machine techniques for automatic leukocyte pattern analysis. *Annals of the New York Academy of Sciences*. 1962; 97: 482-490.
- 164 Prewitt JMS, Mendelsohn ML. The analysis of cell images. *Annals of the New York Academy of Sciences*. 1966; 128: 1035-1053.
- 165 Puente Dominguez JL. Clínica y exploración de los tumores de la mama. *Clin Ginecol*. 1/3. Salvat. Barcelona. 1976.
- 166 Rajeswari S, Chosh SN, Shah PN, Borah VJ. Barr body frequency in the human breast cancer tissue. A prospectivestudy on its prognostic value and its correlation with specific oestradiol receptors. *European Journal of Cancer*. 1977; 13: 99-102.
-



---

167 Remvikos Y, Magdelénat H, Zajdela A. DNA Flow Cytometry applied to fine needle sampling of human breast cancer *Cancer* 61, pp. 1629-1634 1988

168 Rivera-Pomar JM. Métodos morfométricos. Actas del XII Congreso Nacional de la Sociedad Española de Anatomía Patológica. Patología (supl. esp.). 1985: 97-101.

169 Ronstrom L, Tribukait B, Esposti PL. DNA pattern and cytological findings in fine needle aspirates of untreated prostatic tumors. A flow-cytofluorometric study. *Prostate*. 1981; 2: 79-88.

170 Rosenthal DL, McLatchie C, Stern E, Withe BS, Castleman KR. Endocervical columnar cell atypia coincident with cervical neoplasia characterized by digital image analysis. *Acta Cytologica*. 1982; 26: 115-120.

171 Rosner DH, Joy JV, Lane WW. Oral contraceptives and breast cancer prognosis in women aged 35 to 50. *J Surg Oncol*. 1985; 30: 52.

172 Russo J, Frederick J, Ownby, et al. Predictors of recurrence and survival of patients with breast cancer. *American Journal of Clinical Pathology*. 1987; 88: 123-131.

173 Russo J, Russo IH. Biology of disease. Biological and molecular bases of mammary carcinogenesis. *Laboratory Investigation*. 1987; 57: 112-137.

174 Salmon I, Coibion M, Larsimont D, et al. Comparison of fine needle aspirates of breast cancers to imprint smears by means of digital cell image analysis. *Analytical and Quantitative Cytology and Histology*. 1991; 13: 193-200.

175 Sandritter W, Kiefer G, Fiefer R, Salm, R., Moore, G.W., Grimm, H. DNA in heterochromatin. Cytophotometric pattern recognition image analysis among cell nuclei in duct epithelium and in carcinoma of human breast. *Beitr. Path. Bd*. 1974; 151: 87-96.

176 Santos E, Villanueva JR. En: *El Cáncer*. 1ª edición. Ed. Prensa Científica. Barcelona. 1985: 5-8.

177 Savino A, Koss LG. The evaluation of sex chromatin as a prognostic factor in carcinoma of the breast. A preliminar report. *Acta Cytologica*. 1971; 15: 372-374.

178 Schatzkin A, Jones Y, Hoover RN, et al.. Alcohol consumption and breast cancer in the epidemiologic followup study of the first national health and nutrition examination survey. *New England Journal of Medicine*. 1987; 316: 1169.



- 
- 179 Serra J. Image analysis and Mathematical morphology. Academic Press. New York. 1982
- 180 Seshadri R, Matthews C, Dobrovic A, Horstall DJ. The significance of oncogene amplification in primry breast cancer. *Internacional Journal of Cancer*. 1989; 43: 270-272.
- 181 Shabaik AS, Cox CE, Clark RA, Reintge DS, Humphrey EJ, Nicosia SV. Imprint Cytology of needle-localized breast lesions. *Acta Cytologica*. 1993; 37: 10-15.
- 182 Slidders W, Robertson AJ, Livesey AE, Cree IA, Brown RA, Beck JS. A study of the histological basis of textural analysis. *Microsc. Acta* 84: 363-378. 1981
- 183 Smith R, Crocker J. Evaluation of nucleolar organizer region-associated proteins in breast malignancy. *Histopathology*. 1988; 12: 113-125.
- 184 Spyrtos F, Brifford M, Gentile A, Brunet M, Brault C, Desplaces A. Flow cytometric study of DNA distribution in cytopunctures of benign and malignant breast lesions. *Analytical and Quantitative Cytology and Histology*. 1987; 9: 485-494.
- 185 Stenkvist B, Bengtsson E, Eriksson O, Jarkrans T, Nordin B, Westman-Naeser S. Histopathological systems of breast cancer classification: reproducibility and clinical significance. *Journal Clinical Pathology*. 1983; 36: 392-398.
- 186 Stenkvist B, Westman-Naeser S, Vegelius J, et al. Analisis of reproducibility of subjective grading systems for breast carcinoma. *Journal Clinical Pathology*. 1979; 32: 979.
- 187 Stoel BC. Interactive histogram equalization. *Pattern Recognition Letters*. 1990; 11: 247-254.
- 188 Stone M. Cross-validators choice and assesment of statistical predictory. *Journal of the Royal Statistical Society*. 1974; 36: 111-147.
- 189 Strax, P. Valoración selectiva para cáncer de mama. *Clinicas Obstétricas Ginecológicas*. Diciembre 1977.
- 190 Streibl N. Three-dimensional imaging by a microscope. *Journal of the Optical Socciety of America. Serie A*. 1985; 2: 121-127.
- 200 Streibl N. Fundamental restrictions for 3-D light distributions. *Optic*. 1984; 66: 341-354.
- 201 Tandon AK, Clark GM, Chamness GC, Chirgwin JM, McGuire WL. Cathepsin and prognosis in Breast cancer. *The New England Journal of Medicine*. 1990; 322: 297-302.
-



- 
- 202 Takahashi M. Atlas en color de citología del cancer. Barcelona. 1973
- 203 Theillet C, Lidereau R, Escot C, et al. Loss of a c-H-ras-I and aggressive human primary breast carcinomas. Cancer REs. 1986; 46: 4776.
- 204 Thomas DP, Lilientfeld AM. Factores geográficos reproductivos y sociobiológicos. Clinica Ginecológica 3/3. Barcelona. Editorial Salvat. 1978.
- 205 Thor A, Ohuchi N, Horan-Hand P, et al. Biology of disease: ras gene alterations and enhanced levels of ras p21 expresion in a spectrum of benign and malignant human mamary tissues. Laboratory Investigation. 1986; 55: 603.
- 206 UICC (International Union Against Cancer). TNM (Classification of malignant tumours). 3ª ed. Ginebra (Suiza). 1978: 47-54.
- 207 Uriburu JV. La mama. 2ª edición. Buenos Aires. Editorial López. 1977.
- 208 Uytterlinde AM, Schipper NW, Baak JP. Comparison of extend of disease and morphometric and DNA flow cytometric factors in invasive ductal breast cancer. Journal of Clinical Pathology. 1987; 40: 1432-1436.
- 209 Van Bogaert LJ, De Muylder C, Maldague P, Maisin H. Prognostic implications of mean nuclear dismeter in breast cancer. British Journal of Cancer. 1980; 42: 537-541.
- 210 Visilog. User's Guide. 2ª edición. Jouy en Josas (France). Editado por Noesis. 1988.
- 211 Wallgren A, Zajicek J. The prognostic value of the aspiration biopsy smear in mammary carcinoma. Acta Cytologica. 1976; 20: 479-485
- 212 Watanabe S. An automated apparatus for cancer prescreening: CYBEST. Computer Graphics Image Process. 1974; 3: 350-358.
- 213 Weibel ER. Principles and methods for the morphometric study of the lung and other organs. Laboratory Investigation. 1963; 12: 131.
- 214 Wied GL, Bartels PH, Bahr GF, Oldfield DG. Taxonomic Intra-Cellular Analytic System (TICAS) for cell identification. Acta Cytologica. 1968; 12: 180-204.
- 215 Wilson SL, Ehrmann RL. The cytologic diagnosis of breast aspirations. Acta Cytologica. 1978; 22: 470-475.
-



- 
- 216 Willet WC, Stampfer MJ, Colditz GA, et al. Moderate alcohol consumption and the risk of breast cancer. *New England Journal of Medicine*. 1987; 316: 1174.
- 217 Wynder E, Rose D, Cohen L. Diet and breast cancer in causation and therapy. *Cancer*. 1986; 58: 1804-1813.
- 218 Wynder E. Dietary factors related to breast cancer. *Cancer*. 1980; 46 (Suppl.): 899.
- 219 Wynder E, Rose D. Diet and breast cancer. *Hospital Practise*. 1984; 19: 73-88.
- 220 Wittekind C, Schulte E. Computerized morphometric image analysis of cytologic nuclear parameters in breast cancer. *Analytical and Quantitative Cytology and Histology*. 1987; 9: 480-484.
- 221 Wolberg WH, Tanner MA, Loh W-Y, Vanichsetakul N. Statistical approach to fine needle aspiration diagnosis of breast masses. *Acta Cytologica*. 1987; 31: 737-741.
- 222 Wolberg WH, Tanner MA, Loh W-Y. Diagnostic schemes for fine needle aspirates of breast masses. *Analytical and Quantitative Cytology and Histology*. 1988; 10: 225-228.
- 223 Wolberg WH, Tanner MA, Loh W-Y. Fine-needle aspiration breast mass diagnosis. *Archives of Surgery*. 1989; 124: 814-818.
- 224 Wolberg WH, Stret WN, Mangasarian OL. Breast cytology diagnosis with digital image analysis. *Analytical and Quantitative Cytology and Histology*. 1993; 15: 396-404.
- 225 Woley RC, Schreiber K, Koss LG, Karas M, Sherman A. DNA distribution in human colon carcinomas and its relationship to clinical behavior. *Journal of the National Cancer Institute*. 1982; 69: 15-22.
- 226 Zajdela A, De la Riva, LS, Ghossein NA. The relation of prognosis to the nuclear diameter of breast cancer cells obtained by cytologic aspiration. *Acta Cytologica*. 1979; 23: 75-80.
- 227 Zajdela A, Maublanc MA. Aspiration cytology and cytoprognosis of breast lesions. En: *New Frontiers in Mammary Pathology*. Editado por Hollmamm KH, De Brux J, Verley JM. New York. Plenum Press. 1981: 79-97.
- 228 Zetterberg A, Espoti PL. Cytophotometric DNA analysis of aspirated cells from prostatic carcinoma. *Acta Cytologica*. 1976; 20: 46-57.
-



# PROTOCOLO I

## Rasgos del extendido citológico.

### GRUPOS CELULARES

<b>Tamaño</b>	{	<b>Grande</b> <input type="radio"/> <b>Pequeño</b> <input type="radio"/>	<b>Desorden celular</b> ("Amontonamiento")	{	<b>si</b> <input type="radio"/> <b>no</b> <input type="radio"/>
---------------	---	---	---	---	--

### CELULAS SUELTAS SEMEJANTES A LAS DE LOS GRUPOS

<b>Nº de células</b>	<input type="radio"/> +	<input type="radio"/> ++	<input type="radio"/> +++
<b>Células similares</b>	<input type="radio"/> Si	<input type="radio"/> No	
<b>Atípias celulares</b>	<input type="radio"/> +	<input type="radio"/> ++	<input type="radio"/> +++

### NUCLEOS

<b>Tamaño variable</b>	<b>Cromatina</b>	<b>Contorno</b>
<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>
Si      No	Variable      Regular	Regular      Irregular

### NUCLEOLOS

<input type="radio"/> No	{	<b>Si</b> <input type="radio"/>	{	<b>Pequeño</b> <input type="radio"/> <b>Grande</b> <input type="radio"/>
-----------------------------	---	---------------------------------	---	---

### CITOPLASMA

<input type="radio"/> <b>Ausente o escaso</b>	<input type="radio"/> <b>Medio</b>	<input type="radio"/> <b>Abundante</b>
--	---------------------------------------	---

#### Diatésis tumoral

<input type="radio"/>	<input type="radio"/>
Si	No

#### Inflamación

<input type="radio"/>	<input type="radio"/>
Si	No

#### Estroma

<input type="radio"/>	<input type="radio"/>
Si	No



Protocolo II

Caso Num.: <div style="border: 1px solid black; width: 100px; height: 30px; margin-top: 5px;"></div>	Num.Citologia: <hr style="border: 0; border-top: 1px solid black; margin-top: 5px;"/>	Historia Nu
Num.A.P.....:		

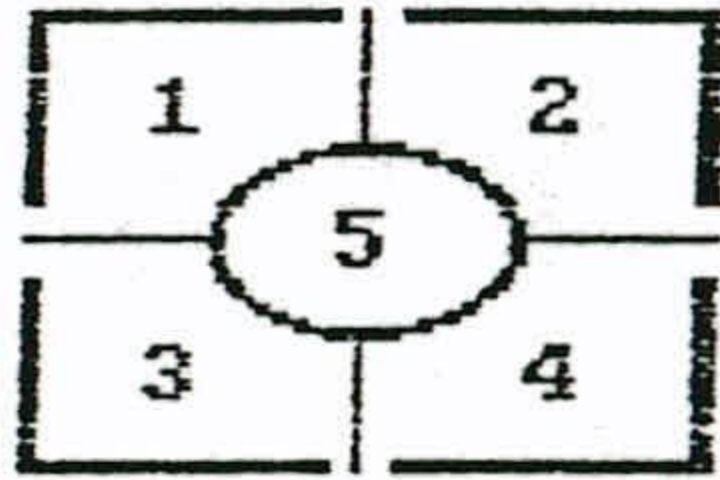
Nombre: \_\_\_\_\_ Edad: \_\_\_\_\_

F.O.: \_ \_ \_ \_ \_

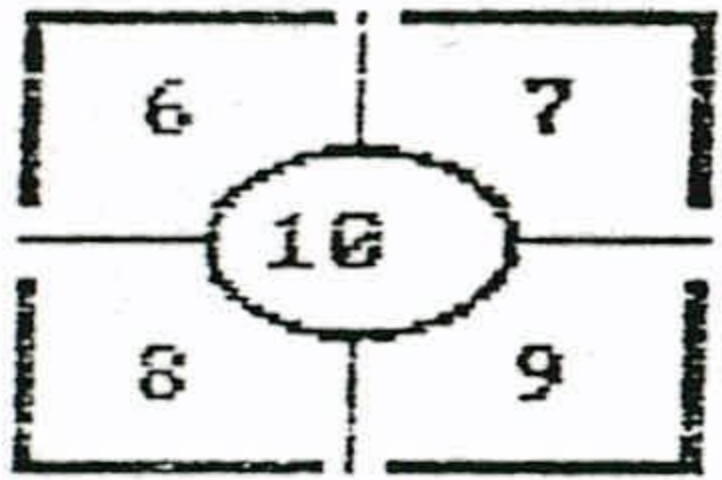
	1	2	3	
EXPLORACION CLINICA _____	+	-	?	1
MAMOGRAFIA _____	+	-	?	2
ECOGRAFIA _____	+	-	?	3
CITOLOGIA _____	+	-	?	4
ANTECEDENTES FAMILIARES _____	+	-	?	5

GRADO DE RIESGO \_\_\_\_\_ TNM: \_\_\_\_\_

LOCALIZACION \_\_\_\_\_



D



I

11: OTROS

\_\_\_\_\_

\_\_\_\_\_

DIAGNOSTICO A.P.: \_\_\_\_\_

FECHA: \_\_\_\_\_

TRATAMIENTO.....: \_\_\_\_\_

EVOLUCION.....: \_\_\_\_\_

FECHA ULTIMA REVISION: \_\_\_\_\_

- Antecedentes F. de cáncer.....
- Primer parto antes de los 30 años.....
- Ultimo parto después de 40 años.....
- Lactancia.....
- Tratamientos hormonales previos.....
- Receptores de estrógenos.....
- " Progesterona.....
- Grupo sanguíneo.....
- Evolución previa.....
- Menarquia.....
- Menopausia.....
- Fumador activo o pasivo.....