



UNIVERSIDAD AUTÓNOMA METROPOLITANA

Casa abierta al tiempo

Azcapotzalco

Maestría en Ciencias de la Computación

Área: Reconocimiento de Patrones

Identificación de tipos de letra

Reporte de Proyecto Terminal II

Presenta:

Juan Villegas Cortez

Matricula: 99380399

Asesor: Dr. Carlos Avilés Cruz

México D.F. Marzo de 2005.

Resumen.

En este estudio nosotros proponemos una alternativa para la necesidad del reconocimiento del tipo de letra, motivo de la realización de este trabajo, basándonos en características de textura globales de las imágenes del documento (asumimos documentos de texto puro, la segmentación de página y el análisis de la estructura del documento están fuera del alcance de este trabajo). No se necesita un análisis local explícito en el método. La clave es el uso del análisis de textura para extraer características globales. Un bloque de texto impreso por cada font puede verse con sus propiedades específicas de textura. La frecuencia espacial y la orientación de contenidos representan las características de cada textura. Son estas características de texturas las usadas para identificar los diferentes fonts.

En principio, cualquier técnica de análisis de textura puede ser aplicada. Nosotros usamos la técnica de aplicar momentos invariantes.

Se basan en el cálculo de los *momentos centrales* como herramienta para la extracción de características propias de las texturas, a su vez, creamos una base de datos de aprendizaje a partir de la imagen de entrada, y aplicamos clasificadores (Gaussiano y KNN). Con esta metodología se obtuvieron como resultados principales dos hechos: el primero, que el nivel de operaciones de cálculo es menor a otros estudios, alcanzando un nivel de acertividad por arriba del 90%, y el segundo, al validar la invariabilidad significativa de los resultados, con este nivel de operaciones, con respecto a la rotación de los textos y su contenido.

El trabajo se organiza del modo siguiente:

- **Capítulo 1.** Introducción.
- **Capítulo 2.** Conceptos teóricos.
- **Capítulo 3.** Detalle del protocolo experimental.
- **Capítulo 4.** Resultados.
- **Capítulo 5.** Conclusiones
- **Capítulo 6.** Perspectivas futuras.
- **Referencias bibliográficas.**

Abstract.

Type-letter recognition is one of the most important aspects in the analysis of documents. The accomplishment of this task is time-consuming and involves a lot of work. Numerous optical recognition techniques have been proposed but only a few of them have taken into account the recognitions of the type of letter or font. Type-letter recognition exerts a lot of influence over the automatic analysis of ADP documents mainly into two aspects: the font affects strongly the recognition of the characters and also the identification of the writing. The font assessment can decrease the number of alternatives in each class turning the recognition of the characters into one kind of font [1]. Also, the output of an ADP document involves not only the content of the document itself but also the font to be used in the printing in order to set only one type of automatic font. In this work we propose an alternative to recognize the font. The study is based in the global-texture characteristics of the images of the document. The following assumptions were made: the document is constituted only by the text, so page segmentation and the analysis of the structure of the document are outside the scope of the study. As a consequence local analysis of the method is not needed. The key aspect of the study is the analysis of the texture which enables the extraction of the global characteristics of the document. A text-block printed with only one font has specific texture properties. These properties comprises the spatial frequency and the orientation of the content of the document. These texture features are used to identify the fonts.

There are several texture-analysis techniques which can be used. It was chosen in this work the Invariant Moment Use technique.

This last technique extracts the texture characteristics through the calculus of the central moments. In addition a data base is created with the input image and Gaussian and KNN classifiers are applied. With the methodology utilized in this work two main results were found. First, the number of the operations performed is lower when compared with other studies; and second, the significant invariability of the results is validated with respect to the rotation of the text and its content. These results reaches a level of confidence above the 90%.

This work is organized as follows.

- **Chapter 1.** Introduction.
- **Chapter 2.** Theoretical concepts.
- **Chapter 3.** Experimental protocol
- **Chapter 4.** Results.
- **Chapter 5.** Conclusions.
- **Chapter 6.** Future Perspectives.
- **References.**

*A mi madre,
a mis hermanas,
a mis sobrinas y sobrino,
y en especial a ti: SCS.*

Tabla de Contenido.

INTRODUCCIÓN.....	9
1.1. ESTADO DEL ARTE DE RECONOCIMIENTO DE TIPO DE LETRA	9
1.2. EL ANÁLISIS DE IMÁGENES.	14
1.2.1. Elementos del análisis de imágenes.....	15
2. CONCEPTOS TEÓRICOS.....	17
2.1. PATRONES Y CLASES DE PATRONES	17
2.2. MÉTODOS DE DECISIÓN TEÓRICA	17
2.2.1. Clasificadores estadísticos óptimos.....	19
2.2.1.1. Fundamentos	19
2.2.1.2. El clasificador de Bayes para clases de patrones gaussianas.	21
2.2.1.3. El clasificador de <i>K</i> -próximos vecinos (KNN).....	25
2.2.1.4. Estimación del error	26
2.2.1.5. Procedimiento de elección del KNN	27
2.2.2. Métodos de evaluación de performance.	29
2.2.2.1. Algoritmo de Conteo de Errores.	29
2.2.2.2. Algoritmo de Resustitución.....	30
2.2.2.3. Algoritmo Holdout (Cross-Validation).	30
2.2.2.4. Algoritmo Leave-One-Out.	31
2.3. EXTRACCIÓN DE ATRIBUTOS DEL FONT.	32
2.3.1. Momentos.	32
3. PROTOCOLO EXPERIMENTAL	36
3.3. ESPECIFICACIONES TÉCNICAS.....	36
3.4. DESARROLLO EXPERIMENTAL	38
3.4.1. El preprocesamiento: la creación de un bloque de texto uniforme.	39
3.4.1.1. Ejemplo de preprocesamiento.	40
3.4.1.2. Detalle del refinamiento del patrón.	45
3.1.1. Extracción de atributos	49
3.1.2. Reconocimiento del tipo de letra	50
4. RESULTADOS EXPERIMENTALES.....	55
4.1. IDENTIFICACIÓN DE TIPO DE FONT.	55
4.2. IDENTIFICACIÓN DE TIPO DE FONT CON ROTACIONES DE: 0°, 45°, 90°, 135°, A 6, 8 10, 12 PUNTOS, PARA EL CASO DE LETRA REGULAR.....	62
4.2.1. Caso: 12 pts, rotación 0°.....	64
4.2.2. Caso: 10 pts, rotación 0°.....	65
4.2.3. Caso: 8 pts, rotación 0°.....	66
4.2.4. Caso: 6 pts, rotación 0°.....	68
4.2.5. Caso: 12 pts, rotación 45°.....	69
4.2.6. Caso: 10 pts, rotación 45°.....	70
4.2.7. Caso: 8 pts, rotación 45°.....	72
4.2.8. Caso: 6 pts, rotación 45°.....	73
4.2.9. Caso: 12 pts, rotación 90°.....	74
4.2.10. Caso: 10 pts, rotación 90°.....	76
4.2.11. Caso: 8 pts, rotación 90°.....	77
4.2.12. Caso: 6 pts, rotación 90°.....	78
4.2.13. Caso: 12 pts, rotación 135°.....	80
4.2.14. Caso: 10 pts, rotación 135°.....	81
4.2.15. Caso: 8 pts, rotación 135°.....	82
4.2.16. Caso: 6 pts, rotación 135°.....	84
4.2.17. Resumiendo.....	85

4.3.	VALIDACIÓN POR CÁLCULO DE MATRIZ DE CONFUSIÓN PARA EL CASO DE LETRA ITÁLICA, NEGRITA Y COMBINACIÓN DE ITÁLICA CON NEGRITA.	88
4.3.1.	<i>Caso: Font a 12 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.</i>	90
4.3.2.	<i>Caso: Font a 10 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.</i>	91
4.3.3.	<i>Caso: Font a 8 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.</i>	92
4.3.4.	<i>Caso: Font a 6 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.</i>	93
4.3.5.	<i>Resumiendo los casos de Itálica, Negrita y combinación de ambas.....</i>	94
4.4.	<i>Experimentos con diferente tamaño de ventana de prueba.</i>	94
4.5.	PRUEBAS CON EL CLASIFICADOR MODIFICADO.	102
4.4.1.	4.5.1. <i>Pruebas sobre el font Regular.</i>	102
4.4.2.	4.5.2. <i>Pruebas sobre los font Itálica, Negrita y Negrita con Itálica.</i>	108
4.6.	RESULTADOS DE SENSIBILIDAD AL RUIDO.....	117
5.	CONCLUSIONES Y PERSPECTIVAS.....	118
	REFERENCIAS BIBLIOGRÁFICAS.	119

Índice de Figuras.

Figura 1.1. Ilustración de tipos de documentos: a) texto, b) página de periódico, c) página de libro de texto de ingeniería, d) código de programa, e) portada de fax, f) página de anuncios clasificados, g) texto con imágenes, h) página de especificaciones técnicas en manual, i) diagrama de operación de máquina, j) diagrama de circuitos electrónicos-eléctricos, k) mapas geográficos, l) listado de valores de acciones en de una casa de bolsa de valores.	10
Figura 1.2. Esquema general para un sistema de identificación de font.	13
Figura 1.3. Elementos del análisis de imágenes.	15
Figura 2.1. Ejemplo de clasificación para tres clases.	18
Figura 2.2. Funciones de densidad de probabilidad para dos clases de patrones unidimensionales. El punto x_0 es la frontera de decisión si las dos clases son equiprobables.	22
Figura 2.3. Selección de los k-vecinos.	27
Figura 2.4. Ejemplo de elección de k-vecinos próximos para un ejercicio de 2 clases en 2 dimensiones en Matlab. El punto a clasificar esta marcado con un círculo.	28
Figura 2.5 Ejemplo de elección de k-vecinos próximos para un ejercicio de 3 clases en 3 dimensiones en Matlab. El punto a clasificar esta marcado con un círculo.	28
Figura 2.6. Estructura de la matriz de confusión.	31
Figura 3.1. Diagrama de bloques de procesamiento de la imagen.	38
Figura 3.2. Imagen de entrada del documento (a), y la proyección horizontal del mismo (b).	40
Figura 3.3. La imagen del documento tras la eliminación de espacios en blanco entre líneas.	41
Figura 3.4. Ejemplo de cálculo de proyección vertical de líneas de texto, solo se ejemplifica con las primeras 7 líneas del texto de la hoja (a), y se muestra la gráfica de sus proyecciones verticales para todas las líneas del documento (b).	42
Figura 3.5. Documento preprocesado eliminando espacios en blanco entre líneas y entre caracteres.	43
Figura 3.6. Documento finalmente preprocesado.	44
Figura 3.7. Análisis de efecto “serif” sobre la estructura del texto.	45
Figura 3.8. Ejemplo de rellenado justo, pero desproporcionado.	46
Figura 3.9. Ejemplos de cuadros resultantes ya procesados.	47
Figura 3.10. Preprocesamiento con corte en bruto de línea de texto rellenada.	48
Figura 3.11. Corte de ventana de muestreo para cálculo del clasificador.	49
Figura 3.1.2. Detalle ilustrativo de extracción de atributos de la imagen: (a) partiendo de la imagen previamente preprocesada, extraemos ventanas barriendo sobre la imagen al azar, y de cada una de estas ventanas de prueba (b), extraemos sus siete momentos invariantes (c)	50
Figura 3.13. (a) Clases generadas con ruido al azar y (b) la imagen muestra que se construye a partir de las clases generadas.	51
Figura 3.14. Imagen construida con las 8 clases para hacer pruebas de reconocimiento del clasificador.	52
Figura 3.15. Procedimiento de Leave-One-Out.	53
Figura 3.16. Procedimiento de Cross-Validation.	54
Figura 4.1. Matriz de confusión con clasificador Gaussiano, con logaritmo.	55

Figura 4.2. Matriz de confusión con clasificador KNN, con logaritmo.	56
Figura 4.3. Matriz de confusión con clasificador Gaussiano, sin logaritmo.	56
Figura 4.4. Matriz de confusión con clasificador KNN, sin logaritmo.	56
Figura 4.5. Matriz de confusión con clasificador Gaussiano.	57
Figura 4.6. Matriz de confusión con clasificador KNN.	57
Figura 4.7. Matriz de confusión con clasificador Gaussiano.	58
Figura 4.8. Matriz de confusión con clasificador KNN.	58
Figura 4.9. Matriz de confusión con clasificador Gaussiano.	58
Figura 4.10. Matriz de confusión con clasificador KNN.	59
Figura 4.11. Matriz de confusión con clasificador Gaussiano.	59
Figura 4.12. Matriz de confusión con clasificador KNN.	59
Figura 4.13. Matriz de confusión con clasificador Gaussiano.	60
Figura 4.14. Matriz de confusión con clasificador KNN.	60
Figura 4.15. Grafica de barras de probabilidades por clasificador y por número de momentos considerados.....	61
Figura 4.16. Muestra de una clase de texto preprocesada sin distorsión, a 0°, 45°, 90° y 135° respectivamente.....	63
Figura 4.17. Muestra de unas clases de texto preprocesada sin distorsión, a 0°, 45°, 90° y 135° respectivamente, con efectos itálica, negrita y combinación de ambas.	88
Figura 4.18. Muestras de ventanas de prueba a 256, 128 y 64 píxeles, de arriba a bajo sucesivamente, con diferentes efectos sobre el font de 10 puntos, a 300 DPI.	101
Figura 4.19. Muestras de textos con ruido aleatorio al 5% (a), 11% (b), 17% (c).	117

Índice de Tablas.

Tabla 4.1. Comparativo de medias de probabilidad con cada clasificador, con y sin logaritmo.....	57
Tabla 4.2. Comparativo de medias de probabilidad con cada clasificador, y con los momentos considerados.....	60
Tabla 4.3. Todos los casos considerados en las pruebas, por cada escala se toman las 4 medidas de rotación (e.g. 12 pts. En 0°, 45°, 90° y 135°).....	62
Tabla 4.4. Resumen de probabilidades medias para el caso de 12 puntos, 0° rotación.	65
Tabla 4.5. Resumen de probabilidades medias para el caso de 10 puntos, 0° rotación.	66
Tabla 4.6. Resumen de probabilidades medias para el caso de 8 puntos, 0° rotación.	67
Tabla 4.7. Resumen de probabilidades medias para el caso de 6 puntos, 0° rotación.	69
Tabla 4.8. Resumen de probabilidades medias para el caso de 12 puntos, 45° rotación.	70
Tabla 4.9. Resumen de probabilidades medias para el caso de 10 puntos, 45° rotación.	71
Tabla 4.10. Resumen de probabilidades medias para el caso de 8 puntos, 45° rotación.	73
Tabla 4.11. Resumen de probabilidades medias para el caso de 6 puntos, 45° rotación.	74
Tabla 4.12. Resumen de probabilidades medias para el caso de 12 puntos, 90° rotación. ..	75
Tabla 4.13. Resumen de probabilidades medias para el caso de 10 puntos, 90° rotación. ..	77
Tabla 4.14. Resumen de probabilidades medias para el caso de 8 puntos, 90° rotación.	78
Tabla 4.15. Resumen de probabilidades medias para el caso de 6 puntos, 90° rotación.	79
Tabla 4.16. Resumen de probabilidades medias para el caso de 12 puntos, 135° rotación.	81
Tabla 4.17. Resumen de probabilidades medias para el caso de 10 puntos, 135° rotación.	82
Tabla 4.18. Resumen de probabilidades medias para el caso de 8 puntos, 135° rotación. ..	83
Tabla 4.19. Resumen de probabilidades medias para el caso de 6 puntos, 135° rotación. ..	85
Tabla 4.21. Todos los casos considerados en las pruebas, por cada escala se toman las 4 medidas de rotación (e.g. 12 pts. En 0°, 45°, 90° y 135°), y 3 efectos (itálica, negrita y combinación de itálica con negrita).....	88
Tabla 4.22. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 12 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.....	90
Tabla 4.23. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 10 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.....	91
Tabla 4.24. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 8 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.....	92
Tabla 4.25. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 6 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.....	93
Tabla 4.26. Medias de probabilidad, con rotación 0°, en font Regular.....	95
Tabla 4.27. Medias de probabilidad, con rotación 0°, en font Itálica.....	96
Tabla 4.28. Medias de probabilidad, con rotación 0°, en font Negrita.....	97
Tabla 4.29. Medias de probabilidad, con rotación 0°, en font Itálica con Negrita.....	98
Tabla 4.30. Medias de probabilidad para el método LOO, con rotación 0°.....	99
Tabla 4.30. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 0°.....	102
Tabla 4.31. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 45°.....	103
Tabla 4.32. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 90°.....	104

Tabla 4.33. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 135°	105
Tabla 4.34. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 64 píxeles, Font Itálica, Rotación 0°, 45°, 90°, 135°	108
Tabla 4.35. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 128 píxeles, Font Itálica, Rotación 0°, 45°, 90°, 135°	109
Tabla 4.36. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 256 píxeles, Font Itálica, Rotación 0°, 45°, 90°, 135°	110
Tabla 4.37. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 64 píxeles, Font Negrita, Rotación 0°, 45°, 90°, 135°	111
Tabla 4.38. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 128 píxeles, Font Negrita, Rotación 0°, 45°, 90°, 135°	112
Tabla 4.39. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 256 píxeles, Font Negrita, Rotación 0°, 45°, 90°, 135°	113
Tabla 4.40. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 64 píxeles, Font Itálica + Negrita, Rotación 0°, 45°, 90°, 135°	114
Tabla 4.41. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 128 píxeles, Font Itálica + Negrita, Rotación 0°, 45°, 90°, 135°	115
Tabla 4.42. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 256 píxeles, Font Itálica + Negrita, Rotación 0°, 45°, 90°, 135°	116

Índice de Gráficas.

Gráfica 4.1.	86
Gráfica 4.2.	86
Gráfica 4.3.	87
Gráfica 4.4.	87
Gráfica 4.5.	90
Gráfica 4.6.	91
Gráfica 4.7.	92
Gráfica 4.8.	93
Gráfica 4.9.	94
Gráfica 4.10.	95
Gráfica 4.11.	96
Gráfica 4.12.	97
Gráfica 4.13.	98
Gráfica 4.14. Los datos de la Tabla 4.30.	100
Gráfica 4.15. De los datos de la Tabla 4.30, solo las medias de probabilidad para el clasificador Gaussiano, por el método de validación de LOO.	100
Gráfica 4.16. De los datos de la Tabla 4.30, solo las medias de probabilidad para el clasificador KNN, por el método de validación de LOO.	101
Gráfica 4.17. Medias probabilidades, clasificador modificado, Font regular, Rotación 0°.	102
Gráfica 4.18. Medias probabilidades, clasificador modificado, Font regular, Rotación 45°.	103
Gráfica 4.19. Medias probabilidades, clasificador modificado, Font regular, Rotación 90°.	104
Gráfica 4.20. Medias probabilidades, clasificador modificado, Font regular, Rotación 135°.	105
Gráfica 4.21. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones de 0°, 45°, 90° y 135°, font Regular.	106
Gráfica 4.22. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones de 0°, 45°, 90° y 135°, font Regular.	106
Gráfica 4.23. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones de 0°, 45°, 90° y 135°, font Regular.	106
Gráfica 4.24. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones 0°, 45°, 90° y 135°, Font Itálica.	108
Gráfica 4.25. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones 0°, 45°, 90° y 135°, Font Itálica.	109
Gráfica 4.26. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones 0°, 45°, 90° y 135°, Font Itálica.	110
Gráfica 4.27. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones 0°, 45°, 90° y 135°, Font Negrita.	111
Gráfica 4.28. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones 0°, 45°, 90° y 135°, Font Negrita.	112
Gráfica 4.29. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones 0°, 45°, 90° y 135°, Font Negrita.	113

Grafica 4.30. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones 0°, 45°, 90° y 135°, Font Itálica + Negrita.....	114
Grafica 4.31. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones 0°, 45°, 90° y 135°, Font Itálica + Negrita.....	115
Grafica 4.32. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones 0°, 45°, 90° y 135°, Font Itálica + Negrita.....	116
Grafica 4.33. Resultados de pruebas de texto con ruido aleatorio al 5, 11 y 17%.	117

Introducción.

1.1. Estado del Arte de Reconocimiento de Tipo de Letra.

Las computadoras se han vuelto más y más poderosas en el procesamiento de datos con el tiempo; esta característica las ha hecho la herramienta perfecta en tratar de implementar algunas de las capacidades humanas. Dado que la visión es el sentido por el cual nosotros, los humanos, adquirimos la mayor proporción de información que usamos en nuestras actividades diarias, varias investigaciones se han hecho tratando de simular el proceso de visión con computadoras. En particular, la capacidad de entender figuras visuales que no son exactamente como modelos conocidos, pero lo suficientemente cercanos a algunos, es una de las características humanas que podemos hacer mejor que las computadoras. Ejemplos de esta clase de figuras son los símbolos que nosotros los humanos usamos para la comunicación escrita, para tener una mejor idea de esto, la Figura 1.1 nos muestra un ejemplo de una variedad de documentos que enfatiza los requerimientos del análisis digital de documentos [13] relativos a la estructura simbólica, este ejemplo nos muestra lo deseable que resulta extraer información suficiente para indexar cada documento, su tipo de contenido: texto, texto con imágenes, diagramas, gráficas, mapas en general, tablas, o combinación de estos tipos de contenidos. Nuestro trabajo se centra en documentos de textos, como la subimagen (a) de la Figura 1.1, los documentos con contenido combinados (e.g. texto con imágenes, texto con gráficos, etc.) quedan fuera del alcance de nuestro estudio.

El reconocimiento de tipo de letra, o de *font*, es una tarea fundamental en la identificación y el análisis de documentos. Así ha sido una tarea difícil que demanda un gran número de recursos de cómputo. Existen diversas técnicas que han sido propuestas para resolver este problema. Las técnicas de reconocimiento óptico de caracteres (OCR) son comúnmente usadas por desarrolladores de software comercial [13], pero el tipo de font es generalmente descuidado. Sin embargo, las técnicas del procesamiento automático de documentos (ADP) tienen que tomar en cuenta el tipo de font como uno de los dos principales aspectos. El primero, el tipo de carácter, que una vez determinado es generalizado en el reconocimiento del font y la identificación final del documento. Como un resultado del solo uso de este alcance, el número de formas alternativas que cada clase de familia de font puede tener es reducido. Esto claramente nos lleva al reconocimiento de una sola clase de font [13]. El segundo aspecto (ideal) que más debe ser considerado en las técnicas de ADP es la identificación de los tipos de font usado dentro del documento. Así ha sido usualmente ignorado en orden de la importancia que se le ha ido dando.

Existen dos aproximaciones para la identificación de tipo de letra, por un lado están los métodos tipográficos y por otro los métodos globales.

A continuación se citan los trabajos más importantes.

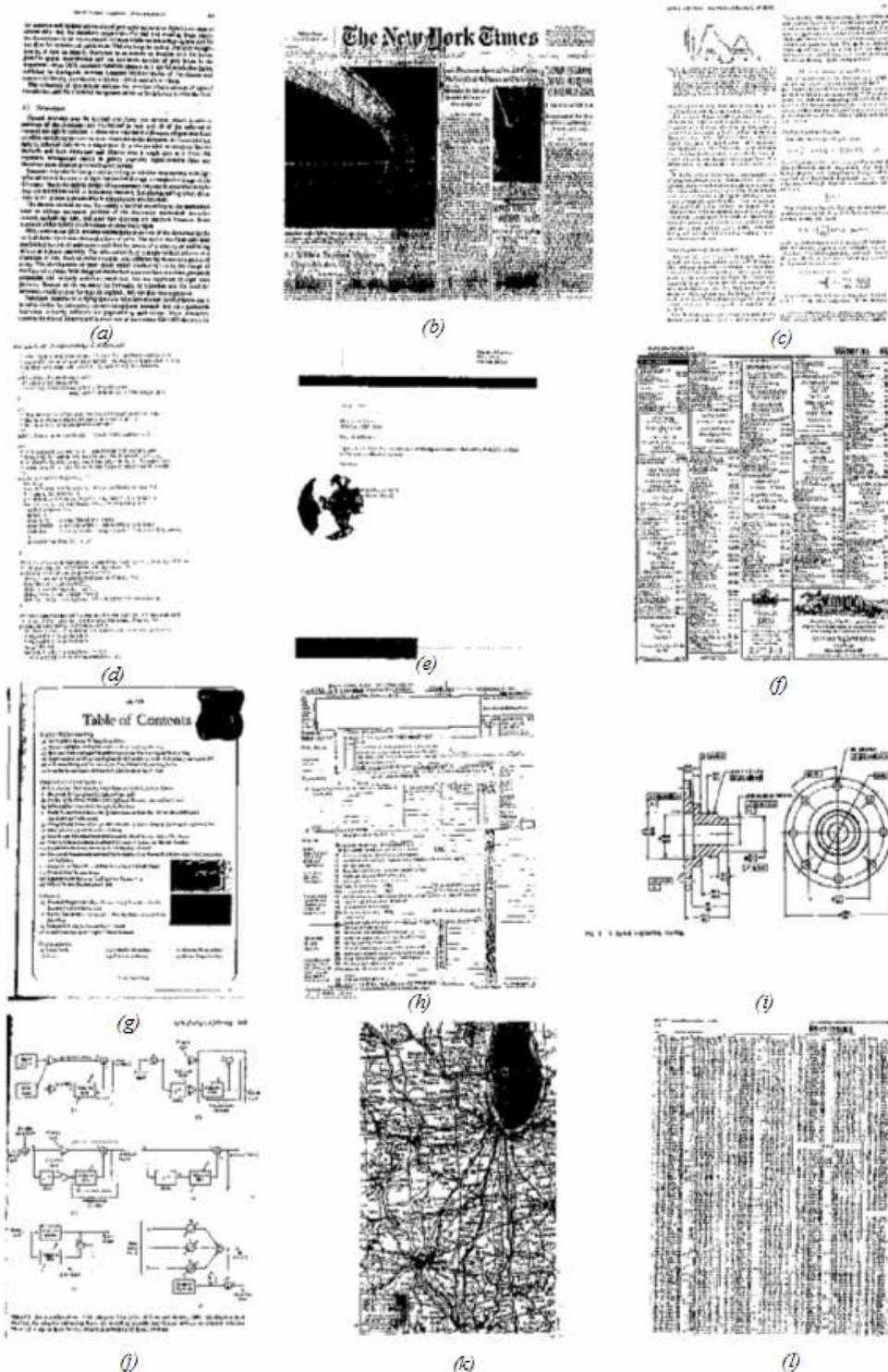


Figura 1.1. Ilustración de tipos de documentos: a) texto, b) página de periódico, c) página de libro de texto de ingeniería, d) código de programa, e) portada de fax, f) página de anuncios clasificados, g) texto con imágenes, h) página de especificaciones técnicas en manual, i) diagrama de operación de máquina, j) diagrama de circuitos electrónicos-eléctricos, k) mapas geográficos, l) listado de valores de acciones en de una casa de bolsa de valores.

Khoubyari y Hull [9] presentan un algoritmo que identifica el font predominante sobre la imagen de un documento impreso. Funciones de palabras frecuentes del idioma inglés (tales como *the, of, and, a, y to*) son también consideradas como parte del reconocimiento del font; tomando racimos de imágenes de palabras, generadas a partir del documento a analizar, y comparándolas con funciones de palabras de una base de datos de funciones de palabras derivadas de imágenes de fonts en documentos. El font base se escoge como el font que es más similar a el que se va a analizar. Este avance en su momento consideró el hecho de que el común de los documentos se realizaba con un solo tipo de font predominantemente, también que las palabras repetidas en un documento son utilizadas en el ruido del mismo.

Cooperman [3] presenta un conjunto de detectores locales que fueron empleados para identificar características individuales de cada font, como la altura, grosor, densidad, líneas, etc., para un sistema de OCR, aquí los autores tratan el tipo de ruido que se presenta para este tipo de reconocimiento, el OCR, identificando atributos del font como el grosor y densidad son los que más cambian en las vecindades de las palabras, mientras que los atributos del *serif* cambian con menos frecuencia dentro de un mismo párrafo, y concluye mostrando que un documento como un todo, tiende a tener un número limitado de atributos.

Shi y Pavlidis [7] hacen reconocimiento de font basados en propiedades de histogramas de palabras, donde propiedades como de inclinación, densidades de histogramas, etc., fueron medidas, haciendo con esto un análisis contextual del documento. La información del documento es extraída desde dos fuentes: una son las propiedades globales de la página del documento, y la otra es el resultado de clasificar palabras cortas en párrafos del documento (del idioma inglés tales como *it, off, etc.*). El procesamiento contextual es hecho por composición de palabras candidatas del resultado de reconocimiento, y entonces se checa cada candidato con un diccionario a través de una revisión ortográfica.

Zramdini e Ingold [20] muestran un avance estadístico para el reconocimiento de font basado en sus aspectos tipográficos globales, para hacer frente al hasta entonces ignorado problema del reconocimiento de font. Realizan la identificación de tipografía, peso, caída y tamaño del texto en un bloque de imagen sin conocimiento alguno del contenido del texto. El reconocimiento se basa en un clasificador Bayesiano multivariado y opera en un conjunto dado de fonts conocidos.

Un avance similar, por Schreyer [15], donde atributos locales de textons¹ fueron usados. Puede verse que todos estos trabajos están basados en aspectos tipográficos que fueron extraídos con análisis locales minuciosos en lugar de análisis global, a excepción del trabajo de [17], mencionado en el párrafo anterior.

Sólo un autor, Zhu [19], se halló que trató de usar análisis de textura global, considerando al documento como una imagen conteniendo características específicas de

¹ Véase Jules y Bergen [8] para una definición de textons, y Malik et al. [10] como un avance alternativo de textons.

texturas, y considerando al reconocimiento de font como una identificación de texturas, independiente del contenido. Los filtros de Gabor fueron implementados a diferentes frecuencias y orientaciones por estos autores, conduciendo los resultados del reconocimiento como una función del ruido pimienta (degradación de reconocimiento como una función de ruido aditivo). Bien que los resultados dados en [16] son buenos, hay espacio para mejoras sustanciales. La experiencia nos muestra que un análisis global nos puede llevar a buenos resultados en el reconocimiento de patrones.

Los autores Aviles-Cruz y Rangel-Kuoppa [2], proponen una nueva técnica para el reconocimiento de font, este avance se basa en el análisis global de texturas, donde los métodos estadísticos son usados para identificar y clasificar las características de los fonts. El reconocimiento del font es realizado tomando al documento como una simple imagen, con uno o varios tipos de fonts presentes, la identificación no es tomada letra a letra como en los avances convencionales, estos autores proponen el uso de barrido de ventanas de pruebas sobre el documento preprocesado² para obtener sus características, usando momentos de tercer y cuarto orden. Esta técnica empleada no hace uso del estudio local de la tipografía, más aún, es independiente del contenido.

Nuestra propuesta de trabajo parte del avance de [2], haciendo uso de características del análisis global de texturas, implementando clasificadores estadísticos, con la particularidad de tomar en cuenta los momentos invariantes de Hue de segundo y tercer orden.

² El preprocesamiento generalmente consiste de una serie de transformaciones imagen-a-imagen. Este no incrementa nuestro conocimiento del contenido del documento, pero puede ayudar a extraerlo [10]. En el Capítulo 3 se ve a detalle el preprocesamiento aplicado en nuestro trabajo.

El trabajo de identificación propuesto en este trabajo se resume en la Figura 1.2. La imagen original es pre-procesada para crear un bloque de texto uniforme, el cual a su vez es usado para extraer sus atributos estadísticos. Se aplica posteriormente un análisis de sus componentes principales para eliminar redundancias lineales, llevándonos a un conjunto reducido de características. Las características se categorizan usando clasificadores Bayesianos, obteniendo el reconocimiento de un font en particular.

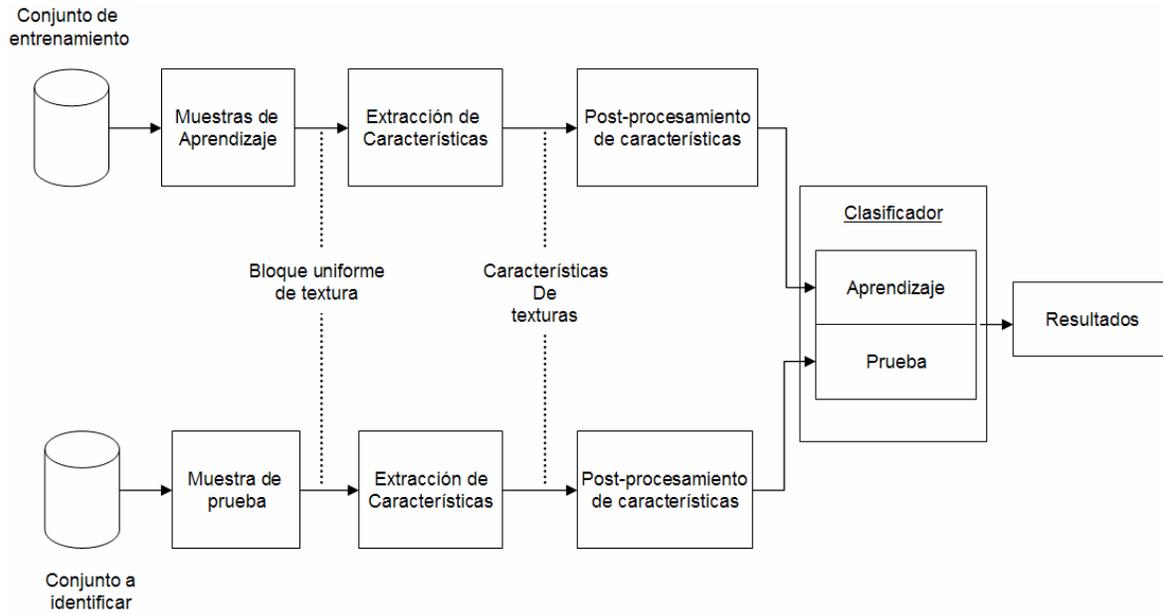


Figura 1.2. Esquema general para un sistema de identificación de font.

1.2. El análisis de imágenes.

El análisis de imágenes es un proceso que consiste en descubrir, identificar y comprender los patrones que son relevantes en el entendimiento de un trabajo basado en imágenes. Uno de los principales objetivos del análisis de imágenes por computadora consiste en dotar a una máquina, en algún sentido, de la capacidad de aproximación, similar a la de los seres humanos. Por ejemplo, en un sistema para la lectura automática de imágenes correspondientes a documentos mecanografiados, los patrones de interés son los patrones alfanuméricos, y el objetivo es conseguir un reconocimiento preciso de caracteres que esté tan próximo como sea posible a la magnífica capacidad exhibida por los seres humanos para realizar tareas de este tipo.

De este modo, un sistema automático de análisis de imágenes debería ser capaz de ofrecer varios grados de inteligencia. El concepto de *inteligencia* es algo vago, particularmente con referencia a una máquina. Sin embargo, no resulta difícil expresar el concepto de los diferentes tipos de comportamiento asociados generalmente con la inteligencia. Diversas características nos vienen rápidamente a la mente:

- 1) La capacidad de extraer la información de interés, separándola de un conjunto de detalles irrelevantes;
- 2) La capacidad de aprender a partir de ejemplos y de generalizar este conocimiento para que se pueda aplicar en nuevas y diferentes circunstancias, y
- 3) La capacidad de realizar inferencias a partir de una información incompleta.

Se puede diseñar e implementar sistemas de análisis de imágenes de estas características para entornos operativos *limitados*. Sin embargo, no sabemos todavía cómo dotar a estos sistemas de un nivel de rendimiento que llegue incluso a emular las capacidades humanas a la hora de realizar funciones genéricas del análisis de imágenes. Gracias a la investigación en sistemas informáticos y biológicos, continuamente se están descubriendo nuevas y prometedoras teorías que intentan explicar el conocimiento visual humano. Sin embargo, las características más avanzadas de análisis de imágenes por computadora se basan en su mayor parte en fórmulas heurísticas, adaptadas para la resolución de problemas específicos. Por ejemplo, algunas máquinas son capaces de leer documentos impresos, con el formato adecuado, a velocidades superiores en varios órdenes de magnitud a las que pueda conseguir el lector más experto. De cualquier forma, los sistemas de este tipo son sumamente especializados y tienen poca o ninguna aplicación en otras tareas. Esto es, la teoría actual y las limitaciones de implementación en el campo del análisis de imágenes aplican soluciones que, en gran medida, dependen del problema.

1.2.1. Elementos del análisis de imágenes

Resulta conceptualmente útil dividir el espectro de técnicas de análisis de imágenes en tres áreas básicas.

1. Procesamiento de bajo nivel.
2. Procesamiento de nivel intermedio.
3. Procesamiento de alto nivel.

Aunque estas subdivisiones no tienen límites definitivos, proporcionan un marco útil para la clasificación de los diversos procesos que son componentes esenciales de un sistema autónomo de análisis de imágenes. La Figura 1.3 muestra estos conceptos. Los bloques de líneas discontinuas se solapan, indicando que no existen límites definidos entre los diferentes procesos. Por ejemplo, la umbralización se puede considerar como una operación de realzado (preprocesamiento) o como una herramienta de segmentación.

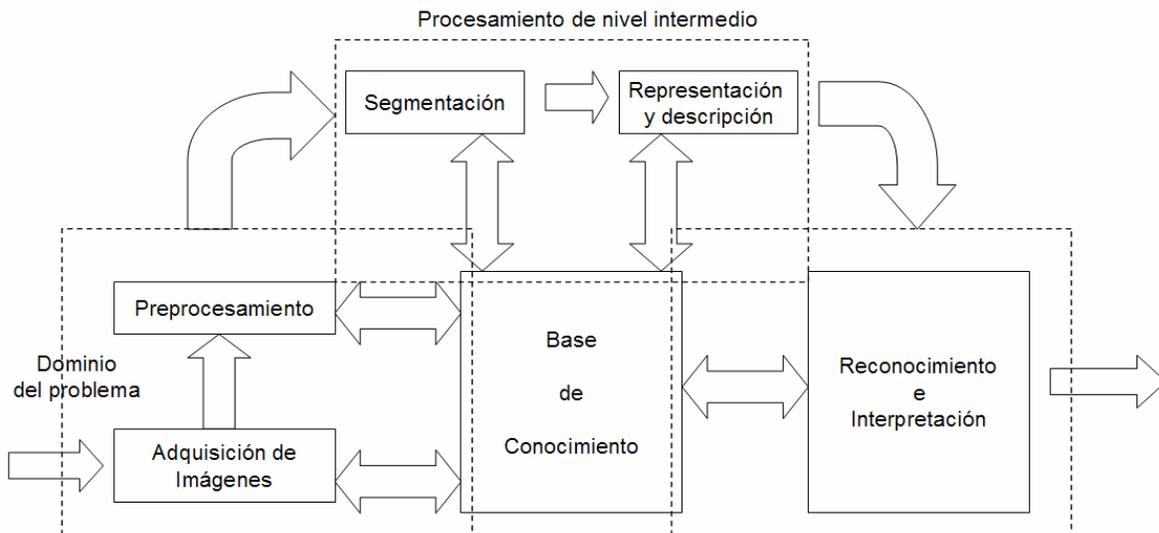


Figura 1.3. Elementos del análisis de imágenes.

En *procesamiento de bajo nivel* trata de un tipo de funciones a las que se puede considerar como reacciones automáticas, y que no requieren inteligencia por parte del sistema de análisis de imágenes. Se considera que la adquisición y el preprocesamiento de imágenes son funciones que operan a bajo nivel. Esta clasificación incluye actividades desde el proceso mismo de formación de las imágenes hasta las diversas compensaciones aplicables, como la reducción del ruido o supresión, en la medida de lo posible, del emborronamiento de las imágenes. Se pueden comparar las funciones de bajo nivel con el proceso de percepción y adaptación que sufre una persona cuando trata de encontrar un asiento en medio de la oscuridad del teatro, tras haber estado expuesta durante un cierto periodo a la claridad del sol. El proceso (inteligente) de encontrar un asiento vacío no puede empezar hasta que esté disponible una imagen apropiada del entorno. El proceso que sigue el cerebro para adaptar el sistema visual de modo que se produzca tal imagen es una reacción inconsciente, automática.

El *procesamiento de nivel intermedio* se refiere a la labor de extracción y caracterización de los componentes (por ejemplo, regiones) de la imagen que se obtiene de un proceso a bajo nivel. Como indica la Figura 1.3, el procesado de nivel intermedio abarca la segmentación y la descripción. Es necesario construir algunas capacidades para el comportamiento inteligente dentro de procedimientos flexibles de segmentación. Por ejemplo, para llenar los pequeños huecos de un contorno segmentado es necesario disponer de elementos más sofisticados para la resolución de problemas que las simples reacciones automáticas de bajo nivel.

Finalmente, el *procesamiento de alto nivel* comprende el reconocimiento y la interpretación. Estos dos procesos tienen un gran parecido con lo que normalmente se conoce como *conocimiento inteligente*. La mayoría de las técnicas empleadas por los procesamientos de nivel bajo e intermedio utiliza un conjunto de formulaciones teóricas razonablemente bien definidas. Sin embargo, como se adelantó en el reconocimiento, y especialmente en la interpretación, nuestro conocimiento y comprensión de los principios fundamentales llega a ser mucho menos preciso y mucho más especulativo. Esta relativa falta de comprensión se traduce al final en una formulación de restricciones e idealizaciones cuyo propósito es reducir la complejidad de las tareas, hasta llegar a un nivel razonable. El producto final es un sistema con capacidades operativas sumamente especializadas.

2. Conceptos Teóricos.

2.1. Patrones y clases de patrones

Como se vio en la sección anterior, en el análisis de una imagen es fundamental realizar, en algún nivel, un reconocimiento de patrones. Aquí, un *patrón* es una descripción estructural o cuantitativa de un objeto o de alguna otra entidad de interés en una imagen. En general, un patrón está formado por uno o más descriptores. En otras palabras, un patrón es una disposición de descriptores, o características. Una *clase de patrones* es una familia de patrones que comparten algunas propiedades comunes. Las clases de patrones se representan por w_1, w_2, \dots, w_M , donde M es el número de clases. El reconocimiento de patrones mediante una máquina supone la utilización de técnicas que permitan asignar los patrones a sus respectivas clases, automáticamente y con la menor intervención humana posible.

Las tres representaciones de patrones principalmente utilizadas en la práctica son los vectores (para descripciones cuantitativas) y las cadenas y los árboles (para descripciones estructurales). Los patrones vectoriales se representan utilizando letras minúsculas en negrita, como x, y , y z , y tienen la forma:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (2.1-1)$$

donde cada componente x_i , representa el i -ésimo descriptor y n es el número de descriptores. Los patrones vectoriales se representan como columnas (esto es, como matrices $n \times 1$). De aquí que un patrón vectorial se puede expresar en la forma mostrada por la ecuación (2.1-1) o en la forma equivalente $x = (x_1, x_2, \dots, x_n)^T$, donde la T indica transposición.

La naturaleza de los componentes de un patrón vectorial x depende de la técnica de medida utilizada para describir el propio patrón físico.

2.2. Métodos de decisión teórica

Los métodos de decisión teórica para el reconocimiento se basan en la utilización de *funciones de decisión* (o *discriminantes*). Supongamos que $x = (x_1, x_2, \dots, x_n)^T$ representa un patrón vectorial de dimensión n . Para M clases de patrones $\omega_1, \omega_2, \dots, \omega_m$, el problema básico en el reconocimiento de patrones por decisión teórica consiste en encontrar M

funciones de decisión $d_1(x), d_2(x), \dots, d_M(x)$, que tengan la propiedad de que, si un patrón x pertenece a la clase ω_i , entonces

$$d_i(x) > d_j(x) \quad j = 1, 2, \dots, M; \quad j \neq i \quad (2.2-1)$$

En otras palabras, un patrón desconocido x pertenece a la i -ésima clase de patrones si, al sustituir la x en todas las funciones de decisión, $d_i(x)$ toma el mayor valor numérico. En caso de empate, los conflictos se resuelven arbitrariamente.

La *frontera de decisión* que separa la clase ω_i de la ω_j viene dada por valores de x para los que $d_i(x) = d_j(x)$ o, expresado de forma equivalente, por los valores de x para los que

$$d_i(x) - d_j(x) = 0 \quad (2.2-2)$$

La práctica común consiste en hacer que la frontera de decisión entre dos clases sea la función $d_{ij}(x) = d_i(x) - d_j(x) = 0$. De esta forma $d_{ij}(x) > 0$, para los patrones de la clase ω_i y $d_{ij}(x) < 0$ para los patrones de la clase ω_j . En la Figura 2.1, se muestra gráficamente para tres clases.

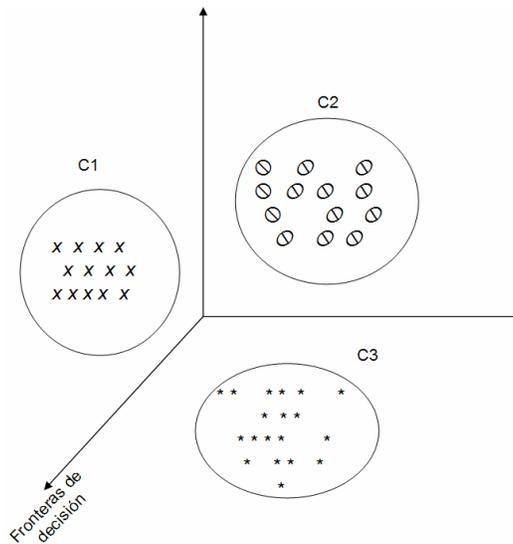


Figura 2.1. Ejemplo de clasificación para tres clases.

2.2.1. Clasificadores estadísticos óptimos

2.2.1.1. Fundamentos

En esta sección desarrollaremos una aproximación probabilística al reconocimiento. Como sucede en la mayoría de los campos relacionados con la medición e interpretación física de acontecimientos, las consideraciones probabilísticas tienen su importancia en el reconocimiento de patrones, debido a la aleatoriedad a la que normalmente está sometida la generación de clases de patrones. Como se muestra en la siguiente exposición, es posible obtener una técnica de clasificación óptima en el sentido de que, por término medio, su empleo tenga asociada la probabilidad más baja de cometer errores de clasificación.

La probabilidad a priori de que un patrón concreto, x , pertenezca de una clase w_i , se representa por $p(w_i/x)$. Si el clasificador de patrones decide que x proviene de w_j cuando realmente proviene de w_i , incurre en un error, de notación L_{ij} . Como el patrón x puede pertenecer a cualquiera de las M clases consideradas, el error medio que se produce al asignar x a la clase w_i es:

$$r_j(x) = \sum_{k=1}^M L_{kj} p(\omega_k/x) \quad (2.2-3)$$

La ecuación (2.2-3) se suele llamar *riesgo medio condicional* o *pérdidas* según la terminología de la teoría de decisiones.

Según la teoría básica de probabilidades, y de acuerdo al teorema de Bayes, la probabilidad condicional se puede expresar como $p(a/b) = [p(a)p(b/a)]/p(b)$. Utilizando esta expresión, escribimos la ecuación (2.2-3) de la forma:

$$r_j(x) = \frac{1}{p(x)} \sum_{k=1}^M L_{kj} p(x/\omega_k) P(\omega_k) \quad (2.2-4)$$

donde $p(x/\omega_k)$ es la función de densidad de probabilidad de los patrones de la clase ω_k y $P(\omega_k)$ es la probabilidad de ocurrencia de la clase w_k . Puesto que $1/p(x)$ es positiva y común a todos los $r_j(x)$, $j = 1, 2, \dots, M$, puede ser eliminada de la ecuación (2.2-4) sin que esto afecte al orden relativo de las funciones, de menor a mayor valor. La expresión del error medio se reduce entonces a

$$r_j(x) = \sum_{k=1}^M L_{kj} p(x/\omega_k) P(\omega_k) \quad (2.2-5)$$

El clasificador tiene M clases posibles para elegir a la hora de clasificar un patrón desconocido dado. Si se calculan $r_1(x), r_2(x), \dots, r_M(x)$ para cada patrón x , y se asigna el

patrón a la clase que tenga el menor error, el error medio total con respecto a todas las decisiones será mínimo. El clasificador que minimiza el error medio total se denomina *Clasificador de Bayes*. Así pues, el clasificador de Bayes asigna un patrón desconocido x a la clase ω_i si $r_i(x) < r_j(x)$ para $j = 1, 2, \dots, M; j \neq i$. En otras palabras, se asigna x a la clase ω_i si

$$\sum_{k=1}^M L_{kj} p(x/\omega_k) < \sum_{q=1}^M L_{qj} p(x/\omega_q) P(\omega_q) \quad (2.2-6)$$

En muchos problemas de reconocimiento, el error de una decisión correcta es cero, y el valor no nulo (por ejemplo, 1) para cualquier decisión incorrecta es siempre el mismo. Bajo estas premisas, la función error para ser

$$L_{ij} = 1 - \delta_{ij} \quad (2.2-7)$$

donde $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$. La ecuación (2.2-7) asigna un error unidad a las decisiones incorrectas y un error cero a las decisiones correctas. Sustituyendo la ecuación (2.2-7) en la (2.2-5) se obtiene:

$$\begin{aligned} r_j(x) &= \sum_{k=1}^M (1 - \delta_{kj}) p(x/\omega_k) P(\omega_k) \\ &= p(x) - p(x/\omega_j) P(\omega_j) \end{aligned} \quad (2.2-8)$$

El clasificador de Bayes asigna entonces un patrón x a la clase w_i si

$$p(x) - p(x/\omega_i) P(\omega_i) < p(x) - p(x/\omega_j) P(\omega_j) \quad (2.2-9)$$

o, de forma equivalente, si

$$p(x/\omega_i) P(\omega_i) < p(x/\omega_j) P(\omega_j) \quad j = 1, 2, \dots, M; j \neq i \quad (2.2-10)$$

Con referencia a la discusión que conduce a la ecuación (2.2-1), vemos que el clasificador de Bayes para funciones de error 0-1 no es nada más que la implementación de las funciones de decisión de la forma

$$d_j(x) = p(x/\omega_j) P(\omega_j) \quad j = 1, 2, \dots, M \quad (2.2-11)$$

donde un patrón vectorial x se asigna a la clase ω_j si $d_i(x) > d_j(x)$ para todo $j \neq i$.

Las funciones de decisión especificadas en la ecuación (2.2-11) son óptimas en tanto que minimizan el error debido a clasificaciones erróneas. Para conseguir esta optimización, sin

embargo, se deben conocer las funciones de densidad de probabilidad de los patrones de cada clase. Este último requisito normalmente no es problema. Por ejemplo, si todas las clases son equiprobables, entonces $P(w_j) = 1/M$. Incluso si esta relación no es cierta, las probabilidades pueden ser inferidas generalmente del conocimiento del problema. La estimación de las funciones de densidad de probabilidad $p(x/w_j)$ es otra cuestión. Si los patrones vectoriales x , son de dimensión n , $p(x/w_j)$ es una función de n variables, que, si su forma no es conocida, requiere métodos de la teoría de probabilidades de varias variables para su estimación. Estos métodos son difíciles de aplicar en la práctica, especialmente si el número de patrones representativos de cada clase no es grande o si las formas de las funciones de densidad de probabilidad no se comportan bien. Por estas razones, el uso del clasificador de Bayes se basa generalmente en la suposición de una expresión analítica para las diversas funciones de densidad y, posteriormente, en una estimación de los parámetros de la expresión para los patrones ejemplo de cada clase. Con diferencia, la forma más comúnmente supuesta para $p(x/w_j)$ es la distribución gaussiana. Cuanto más se aproxime a la realidad esta suposición, más se aproximará el clasificador de Bayes al mínimo error medio de clasificación.

2.2.1.2. El clasificador de Bayes para clases de patrones gaussianas.

Para empezar, consideremos un problema unidimensional ($n = 1$) que incluye dos clases de patrones ($M = 2$) gobernadas por densidades gaussianas, con medias m_1 y m_2 y desviaciones típicas σ_1 y σ_2 , respectivamente. Según la ecuación (2.2-11) las funciones de decisión de Bayes tienen la forma:

$$\begin{aligned}
 d_j(x) &= p(x/w_j)P(\omega_j) \\
 &= \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(x-m_j)^2}{2\sigma_j^2}\right] P(\omega_j) \quad j = 1,2
 \end{aligned}
 \tag{2.2-12}$$

donde los patrones son ahora escalares, de notación x . La Figura 2.2 muestra una representación gráfica de las funciones de densidad de probabilidad para las dos clases. La frontera entre las dos clases es un único punto, denominado x_0 , tal que $d_1(x) = d_2(x)$. Si las dos clases son equiprobables, $P(w_1) = P(w_2) = 1/2$, y la frontera de decisión es el valor x_0 , para el cual $p(x_0/w_1) = p(x_0/w_2)$. Este punto es la intersección de las dos funciones de densidad de probabilidad, como se muestra en la Figura 2.2. Cualquier patrón (punto) situado a la derecha de x_0 se clasifica como perteneciente a la clase w_1 . De manera similar, cualquier patrón situado a la izquierda de x_0 se clasifica como perteneciente a la clase w_2 . Cuando las clases no son equiprobables, x_0 , se mueve a la izquierda si es más probable la ocurrencia de la clase w_1 o, recíprocamente, a la derecha si la clase w_2 tiene más probabilidad de suceder. Este resultado era de esperar, porque el clasificador trata de minimizar el error de clasificación. Por ejemplo, en el caso extremo, si nunca ocurre la

clase w_2 , el clasificador nunca se equivocaría al asignar siempre los patrones a la clase w_1 (esto es, x_0 , se movería hacia menos infinito).

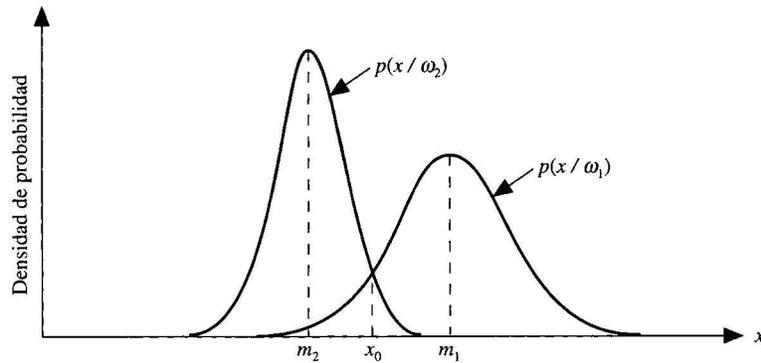


Figura 2.2. Funciones de densidad de probabilidad para dos clases de patrones unidimensionales. El punto x_0 es la frontera de decisión si las dos clases son equiprobables.

En el caso más general de dimensión n , la densidad gaussiana de los vectores de la j -ésima clase de patrones tiene la forma:

$$p(x/\omega_j) = \frac{1}{(2\pi)^{n/2} |C_j|^{1/2}} \exp\left[-\frac{1}{2} (x - m_j)^T C_j^{-1} (x - m_j)\right] \quad (2.2-13)$$

donde cada densidad queda especificada completamente por su vector medio m_j y la matriz de covarianza C_j , que se definen como sigue:

$$m_j = E_j[x] \quad (2.2-14)$$

y

$$C_j = E_j\left[(x - m_j)(x - m_j)^T\right] \quad (2.2-15)$$

donde $E_j[\cdot]$ representa el valor esperado del argumento para los patrones de la clase ω_j . En la ecuación (2.2-13), n es la dimensión de los patrones vectoriales, y $|C_j|$, es el determinante de la matriz C_j . Al aproximar el valor esperado $E_j[\cdot]$ por el valor medio de las cantidades en cuestión se obtiene una estimación del vector medio y la matriz de covarianza:

$$m_j = \frac{1}{N_j} \sum_{x \in \omega_j} x \quad (2.2-16)$$

y

$$C_j = \frac{1}{N_j} \sum_{x \in \omega_j} xx^T - m_j m_j^T \quad (2.2-17)$$

donde N_j es el número de patrones vectoriales de la clase w_j , y la sumatoria se calcula para estos vectores.

La matriz de covarianza C_j es simétrica y no negativa. El elemento C_{kk} , perteneciente a la diagonal de la matriz, es la varianza del k -ésimo elemento de los patrones vectoriales. El elemento C_{jk} , situado fuera de la diagonal, es la covarianza de x_j y x_k . Cuando los elementos x_j y x_k son estadísticamente independientes, $C_{jk} = 0$. La función de densidad de Gauss de varias variables se reduce al producto de la densidad de Gauss (de una variable) de cada elemento de x cuando los elementos fuera de la diagonal de la matriz de covarianza son nulos.

Según la ecuación (2.2-11), la función de decisión de Bayes para la clase ω_j es $d_j(x) = p(x/w_j)P(\omega_j)$. Sin embargo, debido a la forma exponencial de la densidad gaussiana, es más conveniente trabajar con el logaritmo neperiano de la función de decisión. En otras palabras, podemos usar la forma:

$$\begin{aligned} d_j(x) &= \ln[p(x/w_j)P(\omega_j)] \\ &= \ln p(x/w_j) + \ln P(\omega_j) \end{aligned} \quad (2.2-18)$$

Esta expresión es equivalente a la ecuación (2.2-11), en cuanto al rendimiento de la clasificación, puesto que el logaritmo es una función monótona creciente. Dicho de otra forma, el *orden* numérico de las funciones de decisión de las ecuaciones (2.2-11) y (2.2-18) es el mismo. Sustituyendo la ecuación (2.2-13) en la (2.2-18) se obtiene:

$$d_j(x) = \ln P(\omega_j) - \frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |C_j| - \frac{1}{2} [(x - m_j)^T C_j^{-1} (x - m_j)] \quad (2.2-19)$$

El término $(n/2) \ln 2\pi$ es el mismo para todas las clases, por lo que es posible eliminarlo de la ecuación (2.2-19), que pasa a ser:

$$d_j(x) = \ln P(\omega_j) - \frac{1}{2} \ln |C_j| - \frac{1}{2} [(x - m_j)^T C_j^{-1} (x - m_j)] \quad (2.2-20)$$

para $j = 1, 2, \dots, M$. La ecuación (2.2-20) representan las funciones de decisión de Bayes para las clases de patrones gaussianas que tienen una función de error 0-1.

Las funciones de decisión representadas en la ecuación (2.2-20) son hipercuádricas (funciones cuadráticas en el espacio de dimensión n), puesto que en la ecuación no aparece ningún término de grado superior a dos en las componentes de x . Claramente, lo mejor que puede hacer un clasificador de Bayes para patrones gaussianos es situar una superficie general de decisión de segundo orden entre cada par de clases de patrones. No obstante, si las distribuciones de los patrones son realmente gaussianas, ninguna otra superficie conseguirá un error medio inferior en la clasificación.

Si todas las matrices de covarianza son iguales, $C_j = C$, para $j = 1, 2, \dots, M$ –y quitando todos los términos independientes de j – la ecuación (2.2-20) pasa a ser:

$$d_j(x) = \ln P(\omega_j) + x^T C^{-1} m_j - \frac{1}{2} m_j^T C^{-1} m_j \quad (2.2-21)$$

que son funciones de decisión lineales para $j = 1, 2, \dots, M$.

Si, además, $C = I$, siendo I la matriz identidad y $P(\omega_j) = 1/M$ para $j = 1, 2, \dots, M$, entonces

$$d_j(x) = x^T m_j - \frac{1}{2} m_j^T m_j \quad j = 1, 2, \dots, M \quad (2.2-22)$$

La ecuación (2.2-22) representa las funciones de decisión de un clasificador de mínima distancia. Por lo tanto, el clasificador de mínima distancia es óptimo, según Bayes, si: 1) las clases de patrones son gaussianas, 2) todas las matrices de covarianza son iguales a la matriz identidad, y 3) todas las clases son equiprobables. Las clases de patrones gaussianas que satisfacen estas condiciones son nubes esféricas de idéntica forma en n dimensiones (denominadas *hiperesferas*). El clasificador de mínima distancia establece un hiperplano entre cada par de clases, con la propiedad de que el hiperplano es el biselector perpendicular del segmento que une los centros del par de esferas. En dos dimensiones las clases constituyen círculos, y los límites pasan a ser líneas perpendiculares al segmento que une los centros de los círculos, por su punto medio.

Una de las aplicaciones de más éxito del clasificador de Bayes es la clasificación de imágenes obtenidas remotamente, generadas por escáneres multispectrales a bordo de aviones, satélites y estaciones espaciales. En nuestro caso, aplicamos este clasificador al problema de Reconocimiento Óptico de Tipo de Letra, ú *OFR*.

2.2.1.3. El clasificador de K-próximos vecinos (KNN)³.

Este clasificador se basa en la estimación de las K-enésimas distancias, vecindades, de un patrón hacia alguna de las clases en las que se quiere clasificar.

Usando la estimación de clasificación de mínima distancia, dada por $-\ln p_1(X)/p_2(X) \gg t$, donde la escala t es determinada de varias formas correspondiente al tipo de clasificador a trabajar (e.g. Bayes), ahora en esta sección las funciones de densidad son reemplazadas por sus estimadores, que en cuyo caso ahora la estimación del radio para clasificar lo tenemos como

$$-\ln \frac{\hat{p}_1(X)}{\hat{p}_2(X)} = -\ln \frac{\frac{1}{N_1} \sum_{j=1}^{N_1} \kappa_1(X - X_j^{(1)})}{\frac{1}{N_{21}} \sum_{j=1}^{N_2} \kappa_2(X - X_j^{(2)})} \underset{\omega_2}{> t}, \underset{\omega_1}{< t} \quad (2.2-23)$$

con $S = \{X_1^{(1)}, \dots, X_{N_1}^{(1)}, X_1^{(2)}, \dots, X_{N_2}^{(2)}\}$ el conjunto de datos dados. La ecuación (2.2-23) clasifica una muestra de prueba X entre ω_1 o ω_2 , dependiendo si el lado izquierdo de la ecuación es menor o mayor que el estimador t .

Usando el estimador anterior, el radio clasificador ahora es

$$\begin{aligned} -\ln \frac{\hat{p}_1(X)}{\hat{p}_2(X)} &= -\ln \frac{(\kappa_1 - 1)N_2 v_2(X)}{(\kappa_2 - 1)N_1 v_1(X)} \\ &= -n \ln \frac{d_2(X_{\kappa_2 NN}^2, X)}{d_1(X_{\kappa_1 NN}^1, X)} - \ln \frac{(\kappa_1 - 1)N_2 |C_2|^{1/2}}{(\kappa_2 - 1)N_1 |C_1|^{1/2}} \underset{\omega_2}{> t}, \underset{\omega_1}{< t} \end{aligned} \quad (2.2-24)$$

donde de la fórmula del volumen de una esfera, $v = c_0 |A|^{1/2} r^n$, con radio r , en relación a su área de superficie, $v_i = \pi^{n/2} \Gamma^{-1}(n/2 + 1) |C_i|^{1/2} d_i^n$, y $d_i^2(Y, X) = (Y - X)^T C_i^{-1} (Y - X)$. En orden a clasificar una muestra dada X , la k_1 -ésima NN de ω_1 y la k_2 -ésima NN de ω_2 son halladas, las distancias desde X a estas vecindades son medidas, y estas distancias las insertamos dentro de (2.2-24) para probar si es mayor o menor que t . En orden de prevenir complejidades innecesarias asumimos $k_1 = k_2$.

³ De su nombre en idioma inglés: K - Nearest Neighbor

2.2.1.4. Estimación del error

El error de clasificación basado en una muestra de datos dada \mathbf{S} puede ser estimado usando los métodos L y R . Cuando $X_k^{(1)}$ desde ω_1 es probada con el método R , $X_k^{(1)}$ debe ser incluido como un miembro del grupo dado. Por lo tanto, cuando los KNN 's de $X_k^{(1)}$ son encontrados desde ω_1 del grupo dado, $X_k^{(1)}$ por si mismo es incluido entre éstos KNN 's. La Figura 2.3, muestra como los KNN 's son seleccionados y como las distancias a el k -ésimo NN 's son medidas. Note que en la Figura 2.3, el lugar geométrico de puntos equidistantes desde $X_k^{(1)}$ es un elipsoide porque la distancia es normalizada por C_i . También desde $C_1 \neq C_2$ en general, dos diferentes elipsoides son usados para ω_1 y ω_2 . En el método R , $X_k^{(1)}$ y $X_{NN}^{(1)}$ son los próximos y segundos próximos vecinos de $X_k^{(1)}$ desde ω_2 . Así,

$$-\ln \frac{\hat{p}_{1R}(X_k^{(1)})}{\hat{p}_2(X_k^{(1)})} = -n \ln \frac{d_2(X_{2NN}^2, X_k^{(1)})}{d_1(X_{NN}^1, X_k^{(1)})} - \ln \frac{N_2 |C_2|^{1/2}}{N_1 |C_1|^{1/2}} \begin{matrix} \omega_1 \\ > t, < t \\ \omega_2 \end{matrix} \quad (2.2-25)$$

por otro lado, en el método L , $X_k^{(1)}$ ya no es considerado un miembro del grupo dado. Por lo tanto, $X_{NN}^{(1)}$ y $X_{2NN}^{(1)}$ son seleccionados como los próximos y los segundos próximos vecinos de $X_{NN}^{(1)}$ desde ω_1 . La selección de los vecinos de ω_2 es el mismo que en el anterior, así,

$$-\ln \frac{\hat{p}_{1L}(X_k^{(1)})}{\hat{p}_2(X_k^{(1)})} = -n \ln \frac{d_2(X_{2NN}^2, X_k^{(1)})}{d_1(X_{NN}^1, X_k^{(1)})} - \ln \frac{N_2 |C_2|^{1/2}}{N_1 |C_1|^{1/2}} \begin{matrix} \omega_1 \\ > t, < t \\ \omega_2 \end{matrix} \quad (2.2-26)$$

Obviamente, $d_1(X_{2NN}^{(2)}, X_k^{(1)}) \geq d_1(X_{NN}^{(2)}, X_k^{(1)})$, haciendo el lado izquierdo de la ecuación (2.2-26) grande, también lo es para el lado izquierdo de (2.2-25). Así $X_k^{(1)}$ es más probable que sea clasificado incorrectamente en el método L que en el método R .

También hay que notar que en orden encuentra la muestra NN , las distancias a todas las muestras deben ser procesadas y comparadas. Por lo tanto, cuando $d_1(X_{NN}^{(2)}, X_k^{(1)})$ es obtenido, $d_1(X_{2NN}^{(2)}, X_k^{(1)})$ también debe ser obtenido. Estas medidas necesitan tiempo de procesamiento para obtener ambos resultados de L y R que es prácticamente el mismo tiempo que se necesita sólo para el método R .

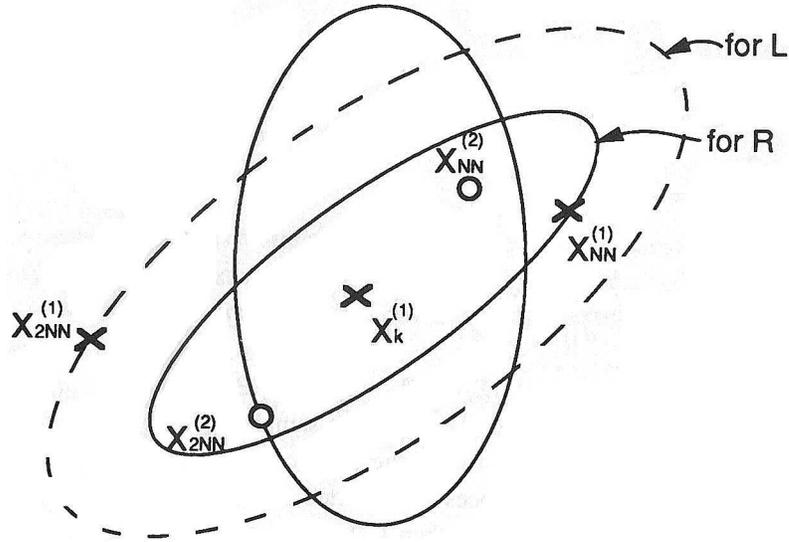


Figura 2.3. Selección de los k-vecinos.

2.2.1.5. Procedimiento de elección del KNN

La aproximación *KNN* mencionada puede ser modificada como se muestra a continuación.

En lugar de seleccionar el k -ésimo *NN* desde cada clase por separado y comparando las distancias, los *KNN*'s de una muestra de prueba son seleccionados desde la mezcla de clases, y el número de vecinos desde cada clase entre las muestras k seleccionadas son contados. La muestra de prueba entonces es clasificada a la clase representada por una mayoría de *KNN*'s. Esto es,

$$\begin{aligned} k_i &= \max \{k_1, \dots, k_L\} \rightarrow X \in \omega_i \\ k_1 + \dots + k_L &= k \end{aligned} \quad (2.2-27)$$

donde k_i es el número de vecinos desde ω_i ($i = 1, \dots, L$) entre los *KNN*'s. Para evitar confusión entre estos dos procedimientos *KNN*, nosotros podemos llamar a la ecuación (2.2-27) *procedimiento de elección del KNN* y a la ecuación (2.2-24) *procedimiento volumétrico KNN*.

Para el procedimiento de elección del *KNN*, es común y práctico usar la misma métrica para medir las distancias de muestras hacia todas las clases, aunque cada clase pudo usar su propia métrica. Así que los k_i 's son enteros y un procedimiento superior es utilizada, es difícil para encontrar un componente de (2.2-27) análogo al umbral de (2.2-24).

Pueden ser mostrados los procedimientos volumétricos *KNN* y de elección $(2k-1)NN$ dando resultados de clasificación idénticos para estas dos clases de problemas usando la misma métrica para ambas clases. Por ejemplo, dejando k y $(2k-1)$ son 3 y 5 respectivamente. En el procedimiento de elección *5NN*, una muestra de prueba es

clasificada a ω_1 , si 2, 4 ó 5 de las 5NN's pertenecen a ω_1 . Esto es equivalente a decir que el 3er NN desde ω_1 es cercano a la muestra de prueba que el 3er NN desde ω_2 .

En la Figura 2.4 mostramos un ejercicio hecho en Matlab, para 2 clases, en 2 dimensiones, en este mismo se pide al usuario ingrese un punto del espacio a clasificar, posteriormente pregunta el número de K-vecinos próximos, y muestra los puntos encerrados en pequeños cuadros en la representación gráfica. El resultado de la corrida se muestra en la misma figura. La Figura 2.5 muestra un ejercicio igual, pero para 3 clases en 3 dimensiones, haciendo referencia a la Figura 2.1, ahora con una visualización concreta.

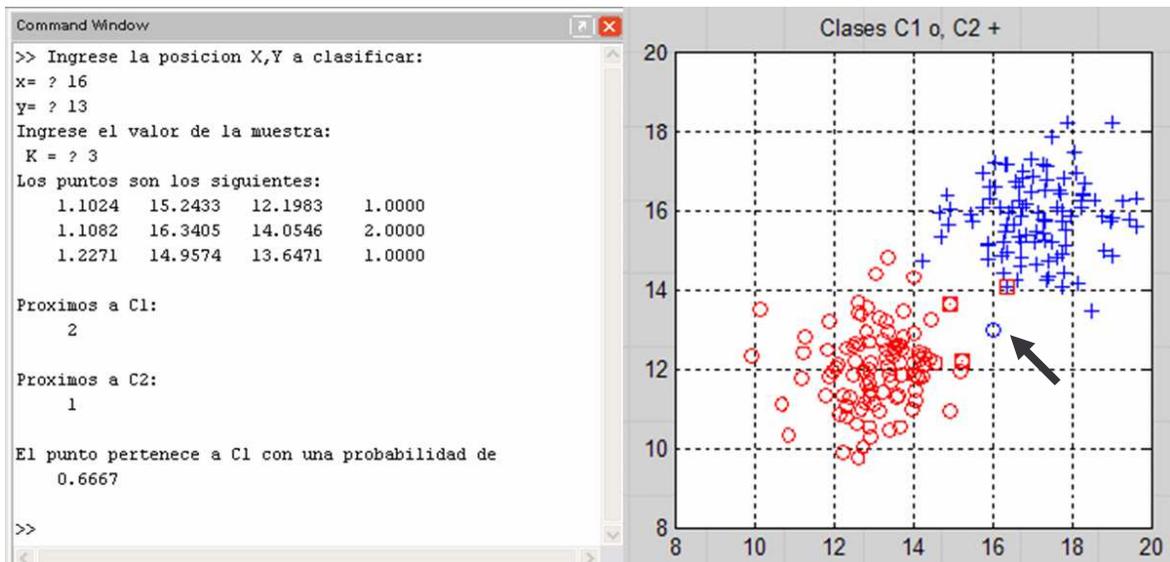


Figura 2.4. Ejemplo de elección de k-vecinos próximos para un ejercicio de 2 clases en 2 dimensiones en Matlab. El punto a clasificar esta marcado con un círculo.

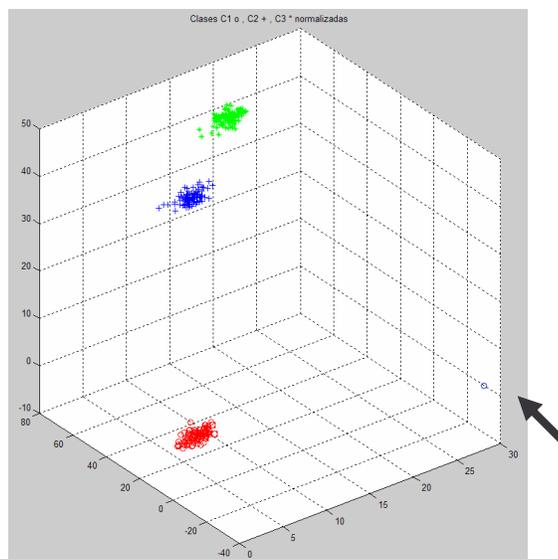


Figura 2.5 Ejemplo de elección de k-vecinos próximos para un ejercicio de 3 clases en 3 dimensiones en Matlab. El punto a clasificar esta marcado con un círculo.

2.2.2. Métodos de evaluación de performance.

Las principales razones de querer una estimación de calidad, es que resulta muy difícil evaluar la calidad de un clasificador por unos sencillos resultados, de saber en que esta configuración es óptima, o ver de si tal clasificador es de mayor calidad que otro.

Definiciones

Tenemos: x un vector a clasificar
 $f_i(x) = 1$ para x clasificado en la clase i
 $f_j(x) = 1$ si no lo es

Se puede definir la matriz de confusión por

$$c_{ij} = \int p(x/w_i) f_j(x) dx$$

con: c_{ij} elemento i, j de la matriz, de tamaño $(ncl \times ncl)$
 ncl número de clases

La matriz de confusión es la manera más difundida de representar simplemente estas probabilidades.

Para evaluar la matriz de confusión, existen diferentes técnicas. Todas las técnicas presentadas aquí pertenecen a la familia de métodos de “conteo de errores”.

2.2.2.1. Algoritmo de Conteo de Errores.

Un error es el hecho de que un elemento del conjunto de aprendizaje perteneciente a la clase i sea clasificado en la clase j .

El conteo de errores es el hecho de evaluar la matriz de confusión contando el número de errores y aciertos de clasificación de un subconjunto de puntos del conjunto de aprendizaje, a través de esto parecerá que las bases dadas de aprendizaje y prueba deben ser estadísticamente independientes (a menos que sean diferentes).

Además, en la práctica, *la cantidad dada disponible es finita y mucho más pequeña que lo que quisiéramos*. Si utilizamos muchas muestras para la clasificación, la cantidad de muestras para la prueba ya no es representativa; y si usamos muchas muestras para la prueba, no deja mucho para el aprendizaje. El clasificador no estará al máximo de sus probabilidades. Hay un compromiso a hacer, es ahí donde intervienen los diferentes algoritmos presentados líneas abajo.

2.2.2.2. Algoritmo de Resustitución.

Algoritmo:

```
 $M_{conf} = ceros(ncl, ncl)$   
aprender(la base entera de las muestras)  
clasificar(esta misma base entera de las muestras)  
conteo de errores
```

La técnica de resustitución es la más simple, la más rápida, pero también la que evalúa menos bien la matriz de confusión: la estimación es demasiado optimista. Intuitivamente, esto puede comprenderse dado que cada elemento a clasificar (del conjunto de prueba) es también un elemento que ha sido aprendido (del conjunto de aprendizaje).

Históricamente, debido al uso del perceptrón, hemos remarcado rápidamente que los resultados buenos de clasificación obtenidos sobre el conjunto de aprendizaje (la resustitución) no eran para nada reproducidos por el conjunto de prueba. Otro ejemplo: una clasificación con un método de las k 's más cercanas en vecindad, y $k=1$ (del vecino más próximo) da un resultado del 100% de clasificación buena sin importar la base de las muestras (!).

2.2.2.3. Algoritmo Holdout (Cross-Validation).

Algoritmo

```
 $M_{conf} = ceros(ncl, ncl)$   
for i=1 to nb_iteraciones(arbitrario)  
  escoger(k elementos al azar dentro de la base de  
  muestras)  
  aprender(toda la base - los elementos k)  
  clasificar(estos k elemento)  
  conteo de errores  
end_for
```

La técnica de Holdout es una respuesta a la falta de desarrollo de la técnica de resustitución, cortamos la base de la muestra en dos partes y aprendemos una y clasificamos la otra. El problema es que el clasificador no está al máximo de sus posibilidades para la clasificación, dado que no ha tomado más que una parte de las muestras. Lo que significa que para k un poco grande, el Holdout tiene una rápida tendencia a subestimar los desarrollos del clasificador y que no utiliza el potencial de sus muestras.

2.2.2.4. Algoritmo Leave-One-Out.

Algoritmo

```

 $M_{conf} = zeros(ncl, ncl)$ 
for i=1 to Nblnd (conjunto de aprendizaje)
    aprender(la base de las muestras menos el elemento j)
    clasificar(el elemento j)
    conteo de error
end_for

```

La técnica del Leave-One-Out es la mejor manera de estimar la matriz de confusión. A cada clasificación, reproducimos las condiciones ideales de clasificación. Clasificamos habiendo aprendido casi la base completa de muestras, (solamente un elemento menos, el que se va a clasificar). Contrariamente al Holdout, no tenemos subestimación de desarrollo. El elemento a clasificar no está dentro de la base de aprendizaje, tampoco tenemos ya el problema de la resustitución.

En contra parte, Leave-One-Out tiene una gran desventaja: pide muchos cálculos. Se necesita tanto aprendizaje como elementos hay.

Como observamos en la figura 2.6, la línea i contiene todos los elementos de la clase i . La primera columna tiene la identificación de las clases. La matriz de confusión proporciona el porcentaje de las clases donde cada clase tiene elementos de la clase i (si están bien clasificados, si son mal clasificados, o si son rechazados). La última columna (opcional) muestra el número de los elementos de la clase i .

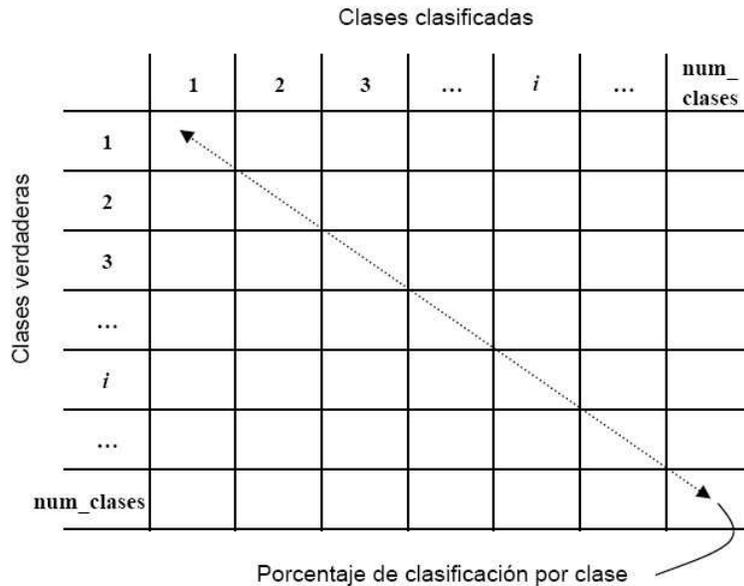


Figura 2.6. Estructura de la matriz de confusión.

2.3. Extracción de atributos del Font.

En teoría hay varios tipos de métodos de análisis de texturas que se pueden usar para la extracción de los atributos del font, uno de ellos es la técnica de los filtros de Gabor [16], otro es la técnica de la matriz de concurrencia de niveles de gris [16]. En los párrafos siguientes describimos la técnica de los *momentos invariantes de Hue*, la que decidimos utilizar para nuestro estudio.

2.3.1. Momentos.

Para una función continua bidimensional $f(x, y)$, el momento de orden $(p+q)$ esta definido por

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (2.2-28)$$

para $p, q = 0, 1, 2, \dots$

El teorema de la unicidad (Papoulis [1965]) establece que si $f(x, y)$ es parcialmente continua y tiene valores diferentes de cero solamente en una parte finita del plano $x y$, existen momentos de todos los ordenes y la secuencia de momentos (m_{pq}) está determinada especialmente por $f(x, y)$. A la inversa, (m_{pq}) está determina especialmente $f(x, y)$. Los *momentos centrales* se pueden expresar como:

$$\mu_{pq} = \int \int (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (2.2-29)$$

donde

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

Para una imagen digital, la ecuación (2.2-29) se convierte en

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (2.2-30)$$

Los momentos centrales de hasta orden 3 son:

$$\begin{aligned}\mu_{10} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^0 f(x, y) \\ &= m_{10} - \frac{m_{10}}{m_{00}} (m_{00}) \\ &= 0\end{aligned}$$

$$\begin{aligned}\mu_{11} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^1 f(x, y) \\ &= m_{11} - \frac{m_{10} m_{01}}{m_{00}}\end{aligned}$$

$$\begin{aligned}\mu_{20} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^0 f(x, y) \\ &= m_{20} - \frac{2m_{10}^2}{m_{00}} + \frac{m_{10}^2}{m_{00}} = m_{20} - \frac{m_{10}^2}{m_{00}}\end{aligned}$$

$$\begin{aligned}\mu_{02} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^2 f(x, y) \\ &= m_{02} - \frac{2m_{01}^2}{m_{00}}\end{aligned}$$

$$\begin{aligned}\mu_{30} &= \sum_x \sum_y (x - \bar{x})^3 (y - \bar{y})^0 f(x, y) \\ &= m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2 m_{10}\end{aligned}$$

$$\begin{aligned}\mu_{12} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^2 f(x, y) \\ &= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2 m_{10}\end{aligned}$$

$$\begin{aligned}\mu_{21} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^1 f(x, y) \\ &= m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2 m_{01}\end{aligned}$$

$$\begin{aligned}\mu_{03} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^3 f(x, y) \\ &= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2 m_{01}\end{aligned}$$

En resumen:

$$\begin{array}{ll}
\mu_{00} = m_{00} & \mu_{11} = m_{11} - \bar{y}m_{10} \\
\mu_{10} = 0 & \mu_{30} = m_{30} - 3\bar{x}m_{20} + 2m_{10}\bar{x}^2 \\
\mu_{01} = 0 & \mu_{12} = m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10} \\
\mu_{20} = m_{20} - \bar{x}m_{10} & \mu_{21} = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01} \\
\mu_{02} = m_{02} - \bar{y}m_{01} & \mu_{03} = m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01}
\end{array}$$

Los *momentos centrales normalizados*, representados por η_{pq} , se definen como

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (2.2-30)$$

donde

$$\gamma = \frac{p+q}{2} + 1 \quad (2.2-31)$$

para $p + q = 2, 3, \dots$

De los momentos segundo y tercero se puede derivar un conjunto de siete *momentos invariantes*:

$$\left. \begin{array}{l}
\phi_1 = \eta_{20} + \eta_{02} \\
\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 = (\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\
\quad - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\
\quad \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \\
\phi_6 = (\eta_{20} - \eta_{02}) \left[(\eta_{30} + 3\eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \\
\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 = (\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\
\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right]
\end{array} \right\} \quad (2.2-32)$$

Este conjunto de momentos es invariante a la traslación, la rotación y al cambio de escala [4], es así que al calcularse para una imagen dada, al rotarla, trasladarla, o cambiándole la escala, sus *momentos* no presentan cambios significativos, y normalmente las pequeñas diferencias que pueden presentar, los errores, se pueden atribuir a la naturaleza digital de los datos.

3. Protocolo experimental

En esta sección presentamos las diferentes etapas de experimentación y desarrollo de este proyecto; así como los requerimientos de hardware y software para la realización del mismo.

3.3. Especificaciones técnicas

Este proyecto fue realizado en Matlab versión 6.5, Release 13, con uso de los siguientes Toolboxes:

- Image Processing Toolbox
- Statistics Toolbox
- Signal Processing Toolbox

Se eligió Matlab dado que ofrece un rápido desarrollo de las aplicaciones, con una interfase amigable y nos permite tener un control rápido sobre los valores de las variables de múltiples programas, además de su valor agregado que consiste en traer incorporadas funciones matemáticas en base a calculo matricial que nos acelera la elaboración de los programas, y por ende no se tienen que desarrollar librerías por separado que hagan cálculos de funciones matemáticas, ya que con otro tipo de software se tendrían que haber sido desarrolladas.

Matlab tiene la facilidad de ser multiplataforma, por lo que se facilita la experimentación directa de las aplicaciones en otros sistemas operativos y otras arquitecturas (32-Bits, 64-Bits / MS-Windows, Unix-Linux, Irix, MacOS).

Para desarrollar y probar nuestro desarrollo tomamos dos tipos de computadoras personales y una tipo servidor, con las siguientes características:

-  Procesador Intel-Pentium IV / 2.80 GHz.
-  512 Mb. RAM
-  Sistema operativo MS-Windows XP Prof.
-  Disco duro de 74.4 Gb.

-  Procesador AMD-Athlon XP / 2.4+ GHz.
-  480 Mb. RAM
-  Sistema operativo MS-Windows XP Prof.
-  Disco duro de 74.4 Gb.

SGI O2

-  Procesador 270 Mhz. MIPS RI2000 (IP32) Procesador con MIPS RI2010 FPU.
-  256 Mb. de memoria RAM.
-  Sistema Operativo IRIX 6.5
-  9 Gb. de disco duro.

La segunda PC (o) se tomo en experimentación para considerar rendimiento con el microprocesador AMD-Athlon.

Para el proceso de digitalización de documentos en papel nos auxiliamos de un scanner común de oficina, en nuestro caso tomamos el modelo:

- HP OfficeJet G85, con el controlador propio bajo sistema operativo MS-Windows XP Profesional, con opción personalizada de digitalización.

Los archivos con código Matlab son guardados con extensión “*.m”, éstos aparecen al final en los anexos de listados de programas.

Todas las imágenes involucradas se tomaron en el formato de mapa de bits, “.BMP”, por la facilidad de manejo del formato, por ser un formato nativo de manejo de imágenes, aunque Matlab nos permite el manejo en otros formatos (JPEG, HDF, PCX, TIFF, etc.); pero haciendo pruebas con las imágenes en formato JPEG, se mostró retardo en el procesamiento por el hecho de que la aplicación tiene que aplicar internamente el proceso de descompresión de la imagen, a cambio, al pasársela en formato BMP obtuvimos un menor tiempo de preprocesamiento.

Estas imágenes están en formato de 8 bits, a escala de grises, y se ingresan a una resolución de 300 DPI, y en dimensiones de hoja de papel tamaño carta (8.1” x 11”), que después del preprocesamiento, descrito posteriormente, quedan en tamaños variables acorde con:

- la cantidad de líneas de texto,
- la longitud de estas, y
- el tipo de font.

La justificación de tomar la imagen a la resolución de 300 DPI, obedece a considerar el estándar de detalle para impresión en papel por litografía digital optima, con esto se tiene un nivel bastante aceptable de detalle del font, es de esperarse que a menor resolución se obtiene un tamaño del documento electrónico menor.

Nos auxiliamos de las siguientes aplicaciones de software para la parte de extracción de las imágenes de los formatos de documentos electrónicos:

- GNU The Gimp (ver. 2.0), para Linux.
- GNU Ghostscript 7.07
- GSview release v4.6
- Jasc Saint Shop Pro. 8
- Driver PostScript de la impresora HP Color LaserJet 4500-PS

3.4. Desarrollo experimental

En la Figura 3.1 presentamos el diagrama de bloques de las tareas desarrolladas por los algoritmos implementados para la realización de este proyecto.

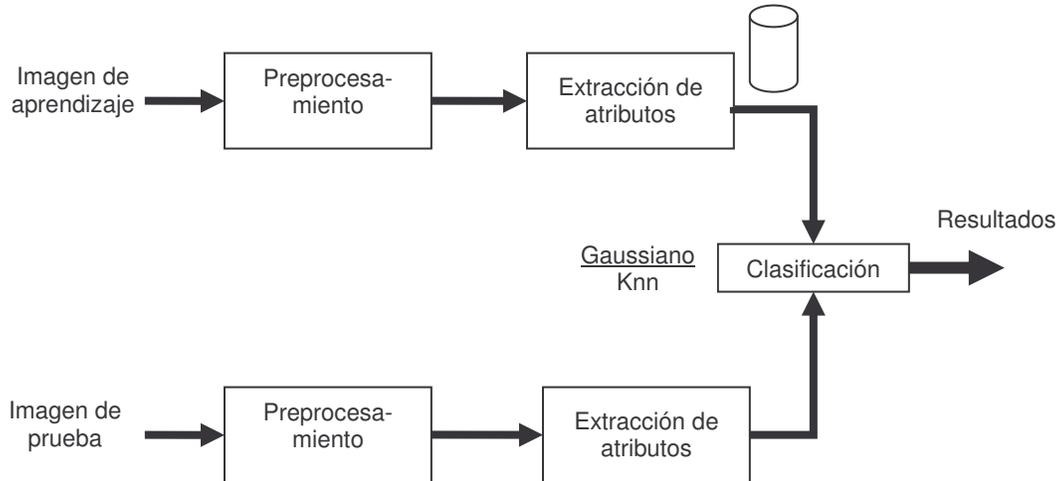


Figura 3.1. Diagrama de bloques de procesamiento de la imagen.

En una primera etapa, la “imagen de aprendizaje” es tomada de una página de texto, elaborada con un procesador de palabras, y se aplican los fonts sobre el mismo texto, para nuestro experimento se consideraron ocho fonts:

1. Courier New
2. Arial
3. Bookman Old Style
4. Franklin Gothic Medium
5. Comic Sans MS
6. Impact
7. Modern
8. Times New Roman

los mismos se etiquetaron con los índices numéricos arriba puestos, y los textos están con tamaño de fuente de 12, 10, 8 y 6 puntos (pts.), las medidas consideradas estándar en los escritos electrónicos comunes.

Cada imagen electrónica se pasa a formato PostScript, para sobre la misma tener libertad de convertirla a la resolución apropiada, en nuestro caso a 300 DPI, y en el formato BMP a escala de grises (8 bits). Aclaramos que Matlab permite también el manejo de imágenes en formato JPEG, este formato tiene la ventaja de comprimir los datos y por lo mismo los archivos generados de las imágenes son de menor tamaño, pero sobre la marcha

de trabajar con muchos archivos de imágenes notamos que aumenta el tiempo de ejecución con las imágenes en este formato, reduciéndose el tiempo si se manejan las imágenes en el formato BMP, que no lleva compresión.

En una segunda etapa, la “imagen de aprendizaje” es tomada de una página de texto con ruido, para hacer pruebas ya no idealizadas.

3.4.1. El preprocesamiento: la creación de un bloque de texto uniforme.

El preprocesamiento de la imagen consiste en aislar el texto:

- Eliminando los espacios en blanco entre líneas de texto (localización de líneas de texto). Para esto calculamos la proyección horizontal del documento (HPP). Los valles entre los picos corresponden a los espacios en blanco entre las líneas de texto. La distancia entre dos valles comprende el ancho vertical de la línea de texto, de esta forma determinamos la altura de cada línea de texto.
- Normalización de línea de texto, se escalan todas las líneas de texto localizadas a un solo tamaño.
- Eliminando los espacios en blanco entre caracteres (normalización de espaciado), esto se desarrolla para reducir influencias no deseables en la textura. Por cada línea de texto se calcula la proyección vertical (VPP). El valle entre los picos de las proyecciones corresponde al espaciado entre caracteres y entre palabras, la distancia entre dos valles corresponde al ancho de cada caracter o palabra, también normalizamos el espaciado escalándolo a un ancho predefinido.
- Rellenando con texto aleatorio entre los espacios en blanco que resultan tras el recorrido de los caracteres o palabras del texto, para lograr un bloque de texto uniforme (*text padding*), podemos usar texto aleatorio para el relleno dado que el método es independiente del contenido.

3.4.1.1. Ejemplo de preprocesamiento.

A continuación mostramos un ejemplo de cómo una de las imágenes del documento electrónico es preprocesada en el orden respectivo, y un detalle a fondo de este proceso se muestra posteriormente.

Cada una de las imágenes de texto se preprocesan en el siguiente orden.

Paso 1. La imagen de la hoja de texto es ingresada, como se muestra en la Figura 3.2 siguiente.

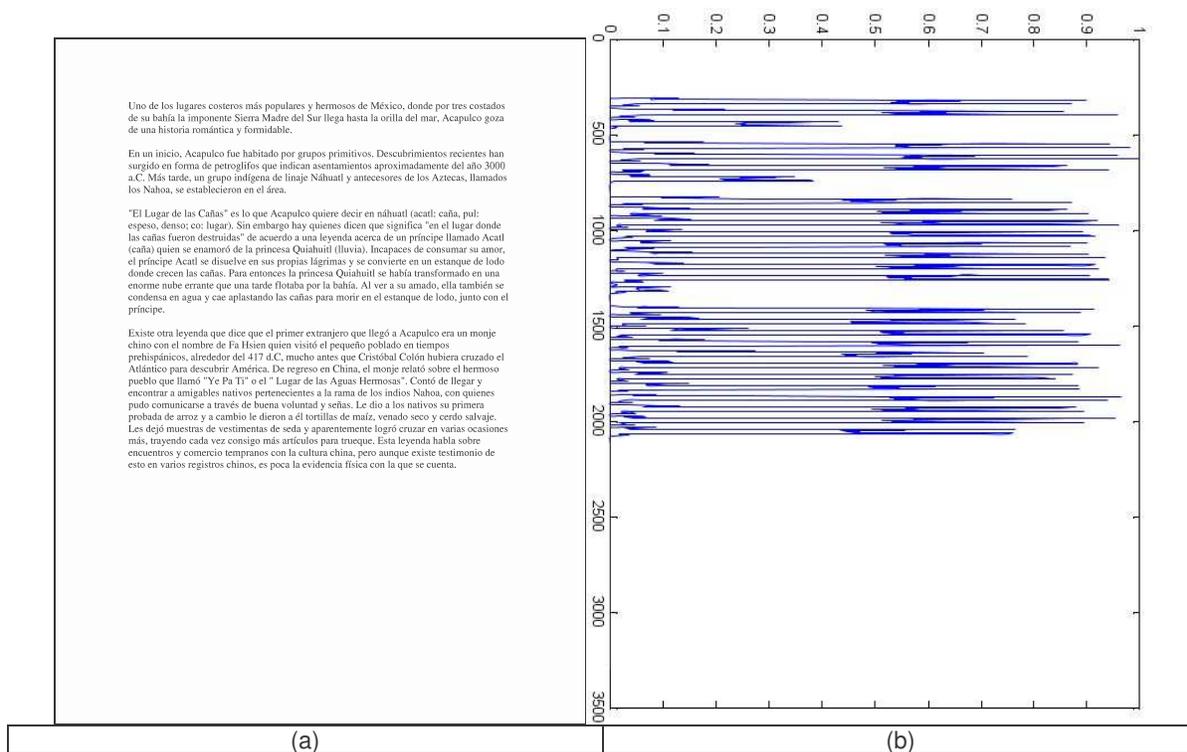


Figura 3.2. Imagen de entrada del documento (a), y la proyección horizontal del mismo (b).

Primero identificamos la existencia de líneas de texto y cuántas de ellas existen en el documento, calculando la proyección horizontal del texto, normalizándola a 1, como se muestra en la imagen anterior.

Paso 2. Auxiliándonos de la proyección horizontal, procedemos a eliminar los espacios en blanco entre líneas de texto del documento, y a aislar cada una de las líneas respectivas, en la Figura 3.3 se muestra el documento sin líneas en blanco intermedias.

Uno de los lugares costeros más populares y hermosos de México, donde por tres costados de su bahía la imponente Sierra Madre del Sur llega hasta la orilla del mar, Acapulco goza de una historia romántica y formidable.

En un inicio, Acapulco fue habitado por grupos primitivos. Descubrimientos recientes han surgido en forma de petroglifos que indican asentamientos aproximadamente del año 3000 a.C. Más tarde, un grupo indígena de linaje Náhuatl y antecesores de los Aztecas, llamados los Nahoas, se establecieron en el área.

"El Lugar de las Cañas" es lo que Acapulco quiere decir en náhuatl (acatl: caña, pul: espeso, denso; co: lugar). Sin embargo hay quienes dicen que significa "en el lugar donde las cañas fueron destruidas" de acuerdo a una leyenda acerca de un príncipe llamado Acatl (caña) quien se enamoró de la princesa Quiahuitl (lluvia). Incapaces de consumar su amor, el príncipe Acatl se disuelve en sus propias lágrimas y se convierte en un estanque de lodo donde crecen las cañas. Para entonces la princesa Quiahuitl se había transformado en una enorme nube errante que una tarde flotaba por la bahía. Al ver a su amado, ella también se condensa en agua y cae aplastando las cañas para morir en el estanque de lodo, junto con el príncipe.

Existe otra leyenda que dice que el primer extranjero que llegó a Acapulco era un monje chino con el nombre de Fa Hsien quien visitó el pequeño poblado en tiempos prehispánicos, alrededor del 417 d.C, mucho antes que Cristóbal Colón hubiera cruzado el Atlántico para descubrir América. De regreso en China, el monje relató sobre el hermoso pueblo que llamó "Ye Pa Ti" o el "Lugar de las Aguas Hermosas". Contó de llegar y encontrar a amigables nativos pertenecientes a la rama de los indios Nahoas, con quienes pudo comunicarse a través de buena voluntad y señas. Le dio a los nativos su primera probada de arroz y a cambio le dieron a él tortillas de maíz, venado seco y cerdo salvaje. Les dejó muestras de vestimentas de seda y aparentemente logró cruzar en varias ocasiones más, trayendo cada vez consigo más artículos para trueque. Esta leyenda habla sobre encuentros y comercio tempranos con la cultura china, pero aunque existe testimonio de esto en varios registros chinos, es poca la evidencia física con la que se cuenta.

Figura 3.3. La imagen del documento tras la eliminación de espacios en blanco entre líneas.

Paso 3. Ahora, sobre cada línea de texto identificada, procedemos a calcular su proyección vertical de texto, como se muestra en las siguientes dos figuras (Figura 3.4. (a) y (b)), en la primera de ella (a) se toma solo una muestra de las primeras líneas, para que se aprecie mejor, y en la segunda imagen (b), se muestra la proyección vertical de todas la líneas identificadas en el documento.

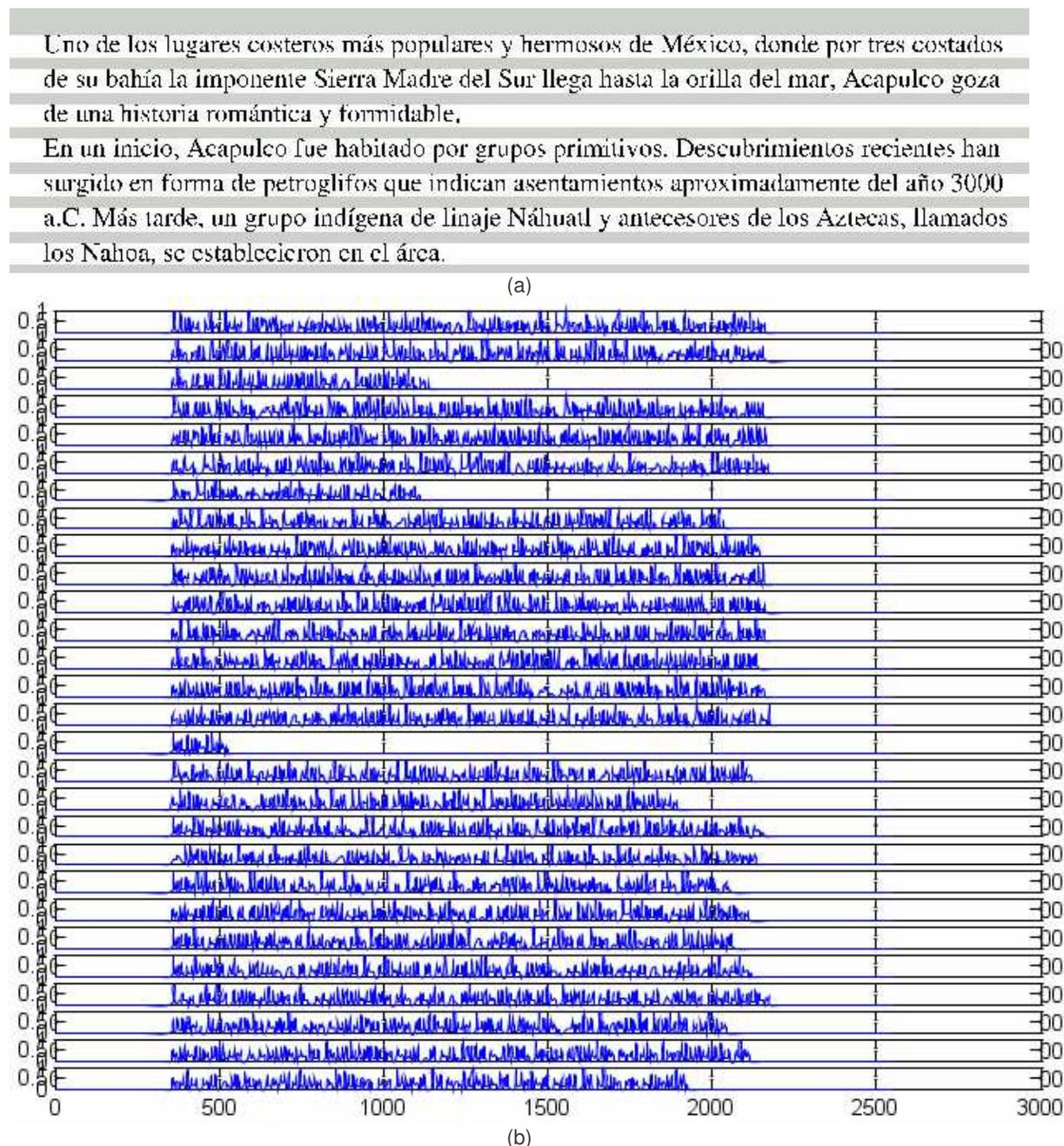


Figura 3.4. Ejemplo de cálculo de proyección vertical de líneas de texto, solo se ejemplifica con las primeras 7 líneas del texto de la hoja (a), y se muestra la gráfica de sus proyecciones verticales para todas las líneas del documento (b).

Paso 4. Procedemos a eliminar los espacios en blanco entre palabras y caracteres, considerando la proyección vertical, y un umbral mínimo, quedando el texto sin blancos como se muestra en la Figura 3.5.

UnodeloslugarescosterosmáspopularesyhermososdeMéxico,doneportrescostados desubahíalaimponenteSierraMadredelSurilegahastalaorilladelmar,Acapulcogoza deunahistoriarománticayformidable.
Enuninicio,Acapulcofuehabitadoporgruposprimitivos.Descubrimientosrecienteshan surgidoenformadepetroglifosqueindicanasentamientosaproximadamentedel año3000 a.C.Mástarde,ungrupoindígenadelinajeNáhuatl yantesoresdelosAztecas,llamados losNahoa,seestablecieronenelárea.
"ElLugardelasCañas"esloqueAcapulcoquiere decir en náhuatl (acatl: caña, pul: espeso, denso, co: lugar). Sin embargo hay quienes dicen que significa "en el lugar donde las cañas fueron destruidas" de acuerdo a una leyenda acerca de un príncipe llamado Acatl (caña) quien se enamoró de la princesa Quiahuitl (lluvia). Incapaces de consumir su amor, el príncipe Acatl se disuelve en sus propias lágrimas y se convierte en un estanque delo donde crecen las cañas. Para entonces la princesa Quiahuitl se había transformado en una enorme nube errante que unatarde flota bapor labahía. Al verasumado, ellatambiense condensa en agua y cae aplastando las cañas para morir en el estanque delo, junto con el príncipe.
Existe otra leyenda que dice que el primer extranjero que llegó a Acapulco era un monje chino con el nombre de Fa Hsien quien visitó el pequeño poblado en tiempos prehispánicos, alrededor del 417 d.C, mucho antes que Cristóbal Colón hubieracruzado el Atlántico para descubrir América. De regreso en China, el monje relató sobre el hermoso pueblo que llamó "Ye Pa Ti" o el "Lugar de las Aguas Hermosas". Contó del legary encontrara amigables nativos pertenecientes a la rama de los indios Nahoa, con quienes pudo comunicarse a través de buena voluntad y señas. Le dio a los nativos su primera probada de arroz y a cambio le dieron a él tortillas de maíz, venado seco y cerdos alvaje. Les dejó muestras de vestimentas de seda y aparentemente logró cruzar en varias ocasiones más, trayendo cada vez consigo más artículos para trueque. Esta leyenda habla sobre encuentros y comercio tempranos con la cultura china, pero aun que existe testimonio de esto en varios registros chinos, es poca la evidencia física con la que se cuenta.

Figura 3.5. Documento procesado eliminando espacios en blanco entre líneas y entre caracteres.

Paso 5. Finalmente aplicamos el Text-Padding para rellenar los espacios en blanco que nos quedan por línea tras el corrimiento de la eliminación de los blancos entre caracteres. La imagen final se muestra en la Figura 3.6, y ésta es ya nuestra imagen preprocesada a considerarse como clase para el muestreo de nuestros clasificadores.

UnodeloslugarescosterosmáspopularesyhermososdeMéxico, dondeportrescostadosUn
 desubahíalaimponenteSierraMadredeelSurlegahastalaorilladelmar, Acapulcogozades
 deunahistoriarománticayformidable.deunahistoriarománticayformidable.deunahisto
 Enuninicio, Acapulcofuehabitadoporgruposprimitivos. DescubrimientosrecienteshanI
 surgidoenformadepetroglifosqueindicanasentamientosaproximadamentedel año3000
 a.C. Mástarde, ungrupoindígenadelinajeNáhuatlyantecesoresdelosAztecas, llamadosa.C
 losNahoa, seestablecieronenelárea. losNahoa, seestablecieronenelárea. losNahoa, seestabl
 "ElLugardelasCañas" esloqueAcapulcoquiere decirennáhuatl (acatl: caña, pul: "ElLugarde
 espeso, denso; co: lugar). Sinembargohayquienesdicenquesignifica"enellugardondeespeso
 lascañasfuerondestruidas" deacuerdoaunaleyendaacercadeunpríncipellamadoAcatillas
 (caña) quienseenamoródelaprincesaQuiahuitl (lluvia). Incapacesdeconsumarsuamor, (ca
 elpríncipeAcatl sedisuelveensuspropiaslágrimasyseconvierteenunestanquedelodoelpr
 dondecrecenlascañas. ParaentonceslaprincesaQuiahuitl sehabíatransformadoenunador
 enormenubeerrantequeunatardeflotabaporlabahía. Alverasuamado, ellatambienseenc
 condensaenaguaycaeaplastandolascañasparamorirenelestanquedelodo, juntoconelcon
 príncipe. príncipe. príncipe. príncipe. príncipe. príncipe. príncipe. príncipe. príncipe. príncip
 ExisteotraleyendaquedicequeelprimerextranjeroquelllegóaAcapulcoeraunmonjeExis
 chinoconelnombredeFaHsienquienvisitóelpequeñopobladoentiemposchinoconelnom
 prehispanicos, alrededordel417d.C, muchoantesqueCristóbalColónhubieracruzadoelpr
 AtlánticoparadescubrirAmérica. DeregresoenChina, elmonjerelatósobreeelhermosoAt
 puebloquellamó"YePaTi" oel"LugardelasAguasHermosas". Contódellegarypuebloquell
 encontraraamigablesnativosperteneientesalaramadelosindiosNahoa, conquienesenco
 pudocomunicarseatravésdebuenavoluntadyseñas. Ledioalosnativossuprimeraipudocor
 probadadearrozyacambioledieronael tortillasdemaíz, venadosecoycerdosalvaje. probad
 LesdejómuestrasdevestimentasdesedayarentementelogrócruzarenvariasocasionesI
 más, trayendocada vez consigomás artículosparatrueque. Estaleyendahablasobremás, tra
 encuentrosycomerciotempranosconlaculturachina, peroaunqueexistetestimoniodeenc
 estoenvariosregistroschinos, espocalaevidenciafísicaconlaquese cuenta. estoenvariosreg

Figura 3.6. Documento finalmente preprocesado.

El tratamiento automatizado del texto nos arrojó un reto a resolver, el detalle de los problemas resueltos se explica en la siguiente sección.

3.4.1.2. Detalle del refinamiento del patrón.

Se trabajó sobre el preprocesamiento del texto ingresado, a 300 DPI de resolución. Se presentaron los siguientes problemas. Considerando la estructura del texto como sigue:

Dado los efectos “serif” de las letras, se tiene una altura de línea de texto variable, una vez que se eliminan los espacios en blanco entre las líneas de texto, esto se ilustra a continuación en la Figura 3.7.

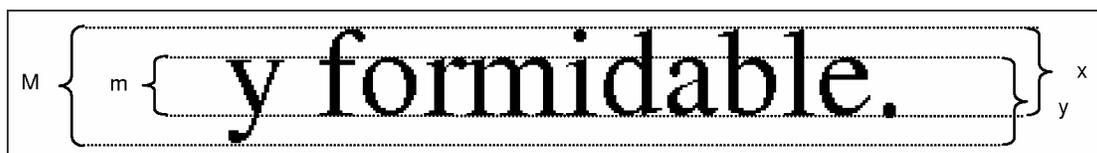


Figura 3.7. Análisis de efecto “serif” sobre la estructura del texto.

De acuerdo con esta ilustración de la Figura 3.7, tenemos que por línea de texto, según el font:

- M*: altura máxima.
- m*: altura mínima.
- x, y*: alturas variables.

Además del caso en que por error del usuario quede una línea de texto conteniendo sólo un punto o una comilla.

Para realizar el *text-padding*, se procedió a calcular por línea de texto el token mínimo, i.e. la mínima palabra o partícula de palabra, en este caso, que se detectará tras un espacio en blanco entre los caracteres, pero como se ve en este línea de ejemplo, para este caso el token sería solo la letra “y”, pero en otros ejemplos puede ser solo letras como “l”, o símbolos como las comillas “”, una sola de ellas, y con estas se procede a rellenar el texto, teniendo líneas finales rellenas no tan buenas para formar un patrón de aprendizaje para nuestro proceso.

Surgieron los siguientes casos:

1. Cuando la línea de texto es pequeña, procedemos a repetirla el número de veces necesario, hasta que la longitud no rebase la línea de texto mayor de todas, a la que se ajustará el ancho del cuadro final ya procesado. Aún así quedará un pequeño espacio sin rellenarse.
2. Se realizó el cálculo del *token mínimo por línea de texto* (TML) y entre todos ellos, se seleccionó el token mínimo global, para con ese tratar de llenar los espacios mínimos que pudieran resultar por llenar; pero aquí nos enfrentamos a el problema de que no siempre el *token mínimo global* (TMG) tiene la misma altura que la línea que se

pretende llenar, entonces no nos sirve. Se trato escalar este token, pero se desproporcione el aspecto de la línea y por ende el patrón global, como se muestra en la Figura 3.7, en la línea de la palabra “área”, así como también se aprecia lo dicho en el punto anterior. Entonces ¿cómo rellenar sin alterar el patrón del font?

Tenemos que si es de la misma altura, el TMG, entonces lo podemos usar para rellenar, pero en caso contrario no, aun después de haber pasado el proceso por el punto anterior.

3. Cuando el espacio en blanco se puede llenar con el mismo token de la línea, pues se usa repitiéndolo hasta no rebasar el tope máximo; pero nos queda aun un espacio en blanco, y entonces volvemos al caso del punto anterior.

Finalmente, tomando en consideración el hecho de no alterar redimensionando el TMG para el relleno faltante, aún así queda un patrón no bueno, o no como lo esperábamos: uniforme en ambas fronteras del cuadro, izquierda y derecha. En la Figura 3.8, presentamos unos cuadros de texto procesados con estas ideas para mejor ilustrar.

Uno de los lugares costeros más populares y hermosos de México, donde por tres U
 costados de subahíal imponente Sierra Madre del Sur legahastala orilla del tete
 mar, Acapulco goza de una historia romántica y formidable. tete tete tete tete tete
 En un inicio, Acapulco fue habitado por grupos primitivos. Descubrimientos tete tete E
 recientes han surgido en forma de petroglifos que indican asentamientos tete tete
 aproximadamente del año 3000 a.C. Más tarde, un grupo indígena de la zona tete tete
 Náhuatl y antecesores de los Aztecas, llamados los Nahoas, se establecieron en el N
 área.
 El Lugar de las Cañas es lo que Acapulco quiere decir en náhuatl (acatl caña, tete tete
 putes peso, denso; cotugar). Sin embargo hay quienes dicen que significa "en el tete p
 lugar donde las cañas fueron destruidas" de acuerdo a una leyenda acerca de un tete
 príncipe llamado Acatl (caña) quien se enamoró de la princesa Quiahuitl (luvia). tete p
 incapaces de consumar su amor, el príncipe Acatl se disuelve en sus propias tete tete
 lágrimas y se convierte en un estanque de lodo donde crecen las cañas. Para tete tete
 entonces la princesa Quiahuitl se había transformado en una enorme nube errante
 que unatarde flotaba por la bahía. Al ver a su amado, ella también se condensó tete
 en agua y cae aplastando las cañas para morir en el estanque de lodo, junto con tete
 el príncipe.
 Existe otra leyenda que dice que el primer extranjero que llegó a Acapulco era un tete
 monje chino con el nombre de Fa-Hsien quien visitó el pequeño poblado en tete tete tete
 tiempos prehispánicos, alrededor del 47 d.C, mucho antes que Cristóbal Colón tete
 hubiera cruzado el Atlántico para descubrir América. De regreso en China, el tete tete
 monje relató sobre el hermoso pueblo que llamó "Ye Pa Ttoe" "Lugar de las tete tete tete
 Aguas Hermosas". Contó de legar y encontrar a amigables nativos pertenecientes
 alarama de los indios Nahoas, con quienes se pudo comunicarse a través de buena te a l
 voluntad y señas. Le dio al nativo su primer aprobado de arroz y a cambio tete tete
 dio a él tortillas de maíz, venado seco y cerdos salvaje. Les dejó muestras de tete tete
 vestimentas de seda y aparentemente le gró cruzar en varias ocasiones más, tete tete
 trayendo cada vez consigo más artículos para trueque. Esta leyenda habla sobre tete
 encuentros y comercio tempranos con la cultura china, pero aunque existetete tete tete
 testimonio de esto en varios registros chinos, es poca la evidencia física con la tete tete
 que se cuenta. tete tete tete

Figura 3.8. Ejemplo de relleno justo, pero desproporcionado.

UnodeloslugarescosterosmáspopularesyhermososdeMéxico, dondeportrestestest
 costadosdesubahíalaimponenteSierraMadredelSurllegahastalaorilladelmar, testtest
 deunahistoriarománticayformidable, testtesttesttesttesttesttesttesttesttest
 Enuninicio, Acapulcofuehabitadoporgruposprimitivos. Descubrimientosrecienteshan
 surgidoenformadepetroglifosqueindicanasantamientosaproximadamentedel año3000
 a.C. Mástarde, ungrupoindígenadelinajeNáhuatl yantecesoresdelosAztecas, test
 llamadoslosNahoa, seestablecieronenelárea.

"ElLugardelasCañas" esloqueAcapulcoquiere decir ennáhuatl (acatl: caña, pul: māmámá
 espeso, denso, colugar). Sinembargohayquienesdicenquesignifica"enellugardonde
 lascañasfuerondestruidas" deacuierdoaunaleyendaacercadeunpríncipellamadoAcatl
 (caña) quienseenamoródelaprincesaQuiahuitl (luvia). Incapacesdeconsumar
 suamor, elpríncipeAcatl sedisuelveensuspropiaslágrimasyseconvierteenunestanque
 de lododondecrecenlascañas. ParaentonceslaprincesaQuiahuitl sehabía
 transformadoenunamonuberrantequeunatardeflota baporlabahía. Alverasuamado,
 ellatambiénsecondensaenaguaycaeplastandolascañasparamori enelam
 estanquedelodo, juntoconelpríncipe. Estanquedelodo, juntoconelpríncipe, estaesta
 Existetraleayendaquedicequeel primerextranjeroquellegó aAcapulcoeraunmonjetest
 chinoconel nombredeFaHsienquienvisitóelpequeñopobladoentiempostesttesttestch
 prehispanicos, alrededordel417d.C, muchoantesqueCristóbalColónhubieracruzadoel
 Atlánticoparadescribir América. Deregreso enChina, elmonjerelatósobreeelhermoso
 puebloquellamó"YePaTi" oel"LugardelasAguasHermosas". Contódellegarytesttesttest
 encontraraamigablesnativosperteneientesalaramadelosindiosNahoa, conquienes
 pudocomunicarseatravésdebuenavoluntadyseñas. Ledioalosnativossu primeratest
 probadade arrozyacambioledieronael tortillasdemaíz, venadosecoycerdosalvaje. prpr
 Lesdejómuestrasdevestimentasdesedayaparentementelogrócruzarenvarias test
 ocasiones más, trayendocada vezconsigomás artículosparatrueque. Estaleyendahabla
 sobreencuentrosycomerciotempranosconlaculturachina, pero aunqueexistesoso
 testimonio deesto envariosregistroschinos, espocala evidenciáfisicaconlaquesetest
 cuenta.cuenta.cuenta.cuenta.cuenta.cuenta.cuenta.cuenta.cuenta.cuenta.cu

UnodeloslugarescosterosmáspopularesyhermososdeMéxico, dondeportrestestest
 costadosdesubahíalaimponenteSierraMadredelSurllegahastalaorilladelmar, Acapulcogozamá
 deunahistoriarománticayformidable, deunahistoriarománticayformidable, māmámá
 Enuninicio, Acapulcofuehabitadoporgruposprimitivos. Descubrimientosrecienteshan
 surgidoenformadepetroglifosqueindicanasantamientosaproximadamentedel año3000
 a.C. Mástarde, ungrupoindígenadelinajeNáhuatl yantecesoresdelosAztecas, llama
 dosalosNahoa, seestablecieronenelárea. lol
 "ElLugardelasCañas" esloqueAcapulcoquiere decir ennáhuatl (acatl: caña, pul: māmámá
 espeso, denso, colugar). Sinembargohayquienesdicenquesignifica"enellugardonde
 lascañasfuerondestruidas" deacuierdoaunaleyendaacercadeunpríncipellamadoAcatl
 (caña) quienseenamoródelaprincesaQuiahuitl (luvia). Incapacesdeconsumar suamor,
 elpríncipeAcatl sedisuelveensuspropiaslágrimasyseconvierteenunestanquedelodo
 dondecrecenlascañas. ParaentonceslaprincesaQuiahuitl sehabía transformadoenunado
 enmonuberrantequeunatardeflota baporlabahía. Alverasuamado, ellatambiénse
 condensaenaguaycaeplastandolascañasparamori enel estanquedelodo, juntoconel má
 príncipe, príncipe, príncipe, príncipe, príncipe, príncipe, príncipe, príncipe, príncipe, prínc
 Existe traleyendaquedicequeel primerextranjeroquellegó aAcapulcoeraunmonjé
 chinoconel nombredeFaHsienquienvisitóelpequeñopobladoentiemposmāmámámá
 prehispanicos, alrededordel417d.C, muchoantesqueCristóbalColónhubieracruzadoel
 Atlánticoparadescribir América. Deregreso enChina, elmonjerelatósobreeelhermoso
 puebloquellamó"YePaTi" oel"LugardelasAguasHermosas". Contódellegarymāmámá
 encontraraamigablesnativosperteneientesalaramadelosindiosNahoa, conquienes má
 pudocomunicarseatravésdebuenavoluntadyseñas. Ledioalosnativossu primeramámá
 probadade arrozyacambioledieronael tortillasdemaíz, venadosecoycerdosalvaje. māmá
 Lesdejómuestrasdevestimentasdesedayaparentementelogrócruzarenvariasocasiones
 más, trayendocada vezconsigomás artículosparatrueque. Estaleyendahabla
 sobreencuentrosycomerciotempranosconlaculturachina, pero aunqueexistetestimonio de
 esto envariosregistroschinos, espocala evidenciáfisicaconlaquesesecuenta. māmámámámá

Figura 3.9. Ejemplos de cuadros resultantes ya procesados.

El programa hecho para este preprocesamiento tiene ajustes para identificar la longitud de la línea de texto, pero estos ajustes traducidos a parámetros numéricos cambian acorde con el tipo de letra, y las sutilezas que se presenten, e.g. en el font "Modern" o "Times New Roman", con palabras como "inca", el punto de la "i" su distancia hacia el inicio de la vertical que define a la letra, cuando delante de la línea no existe otro carácter más alto,

como una "l" o una "t", el programa separa este "punto" de la "i" como una línea de texto más, y con ello altera el procesamiento, ya que también trata de rellenar esa línea.

Finalmente se procedió por un proceso más simple y que nos lleva a tener un patrón más definido para usar en el clasificador, ya que se pretende mejorar la efectividad alcanzada previamente con el preprocesamiento hecho a mano. El proceso es el siguiente:

1. Se calcula la línea de texto, ya sin espacios en blanco, más larga de todas, a esta se ajustaran en largo todas las demás líneas de texto, esta medida es el TOPE.
2. Tomamos una línea de texto, comparamos su longitud, si es igual que TOPE, queda igual la línea preprocesada final.
3. Si no, se repite la línea tantas veces sea necesario hasta apenas medir más que TOPE, y después se trunca en bruto, ajustándola a la longitud de TOPE.
4. Armamos el cuadro de texto preprocesado juntando todas las líneas así procesadas.

En la Figura 3.10 se muestra un ejemplo de una clase ya procesada con el proceso anterior.

UnodeloslugarescosterosmáspopularesyhermososdeMéxico,dondeportrescostadosdeU
subahíala imponenteSierraMadredelSur llegahastalaorilladelmar,Acapulcogozadesu
unahistoriarománticayformidable.unahistoriarománticayformidable.unahistoriaro
Enuninicio,Acapulcofuehabitadoporgruposprimitivos.Descubrimientosrecienteshan
surgidoinformadepetroglifosqueindicanasentamientosaproximadamentedel año3000
a.C.Mástarde,ungrupoindígenadelinajeNáhuatl yantesoresdelosAztecas,llamadosa.C
losNahoa,seestablecieronenelárea.losNahoa,seestablecieronenelárea.losNahoa,seestabl
"ElLugardelasCañas"esloqueAcapulcoquieredecirennáhuatl(acatl:caña,pulespeso,"El
denso;co:lugar).Sinembargohayquienesdicenquesignifica"enellugardondelascañasdens
fuerondestruidas"deacuerdoaunaleyendaacercadeunpríncipellamadoAcatl(caña)fue
quien seenamoródelaprincesaQuiahuitl(lluvia).Incapacesdeconsumarsuamor,elquien
príncipeAcatl se disuelveensuspropiaslágrimasyseconvierteenunestanquedelodoprínc
dondecrecenlascañas.ParaentonceslaprincesaQuiahuitl sehabíatransformadoenunade
enormenubeerrantequeunatardeflotabaporlabahía.Alverasuamado,ellatambienseenc
condensaenaguaycaaplstandolascañasparamorirenelestanquedelodo,juntoconelcon
príncipe.príncipe.príncipe.príncipe.príncipe.príncipe.príncipe.príncipe.príncipe.prínc
ExisteotraleyendaquedicequeelprimerextranjeroquellegóaAcapulcoeraunmonjeExis
chinoconelnombredeFa.Hsienquienvisitóelpequeñopobladoentiemposprehispánicos,c
alrededordel417d.C,muchoantesqueCristóbalColónhubiera cruzadoelAtlánticoparaal
descubrirAmérica.DeregresoenChina,elmonjerelatósobrelhermosopuebloquellamóde
"YePaTi"oel"LugardelasAguasHermosas".Contódellegaryencontraraamigables"YePa
nativospertenecientesalaramadelosindiosNahoa,conquienespudocomunicarseatravés
debuenavoluntadyseñas.Ledioalosnativossuprimera probadadearrozyacambioledebue
dieronaeltortillasdemaíz,venadosecoycerdosalvaje.Lesdejómuestrasdevestimentasdie
desedayaparentementelogró cruzarenvariasocasionesmás,trayendocada vezconsigodese
másartículosparatrueque.Estaleyendahablasobreencontrosycomerciotempranoscom
laculturachina,peroaunqueexistetestimoniodeestoenvariosregistroschinos,especialala
evidenciafísicaconlaquese cuenta.evidenciafísicaconlaquese cuenta.evidenciafísicacon

Figura 3.10. Preprocesamiento con corte en bruto de línea de texto rellena da.

Las justificantes que se tomaron para optar por este camino, después de muchos intentos, fue principalmente el considerar que los clasificadores finalmente para armar su base de datos toman muestreo al azar sobre el cuadro de texto preprocesado, con pequeñas ventanas

de dimensión variable, en nuestro caso se toman cien ventanas en muestreo, y estas mismas ventanas “cortan” el texto en bruto y sobre cada ventana se hace el calculo de los atributos. La Figura 3.11 muestra un corte típico de un cuadro de muestreo.

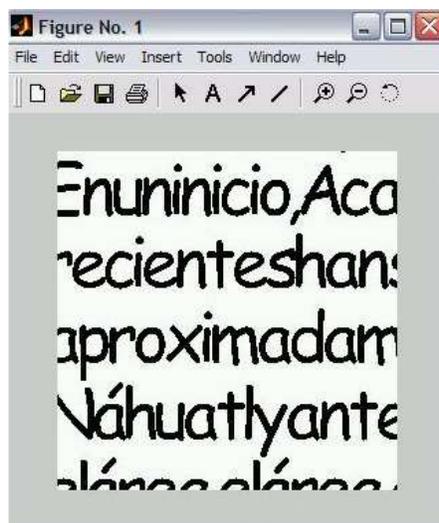


Figura 3.11. Corte de ventana de muestreo para cálculo del clasificador.

Con esta idea obtener los cuadros de texto preprocesados para nuestras ocho clases con las que hacemos la práctica de nuestro procesador se hicieron los experimentos.

3.1.1. Extracción de atributos

Conformamos nuestra base de aprendizaje de la siguiente forma, de cada clase tomamos muestras aleatorias a partir de un barrido de ventanas sobre la imagen del texto (test-window), en un principio de dimensión 128 x 128 píxeles cada ventana, en total tomamos una muestra de cien de estas ventanas, y a continuación a cada una de estas ventanas le aplicamos el cálculo de los siete momentos invariantes, cuya teoría se explicó en el capítulo anterior. De esta forma, por cada ventana obtenemos un vector de siete atributos únicos, sus siete momentos invariantes, y de esta forma cada clase nos queda representada por una matriz de 100 x 7, cien vectores de 7 componentes, y en conjunto con las ocho clases trabajadas tenemos nuestra base de datos de aprendizaje conformada por una matriz única de 800 x 7. Como se muestra en la siguiente Figura 3.12.

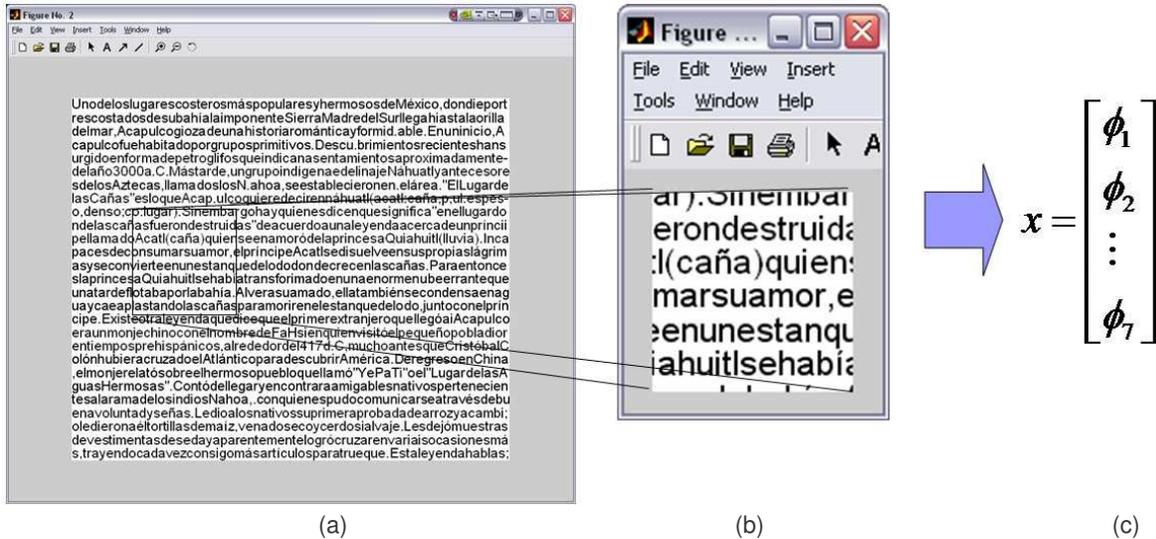


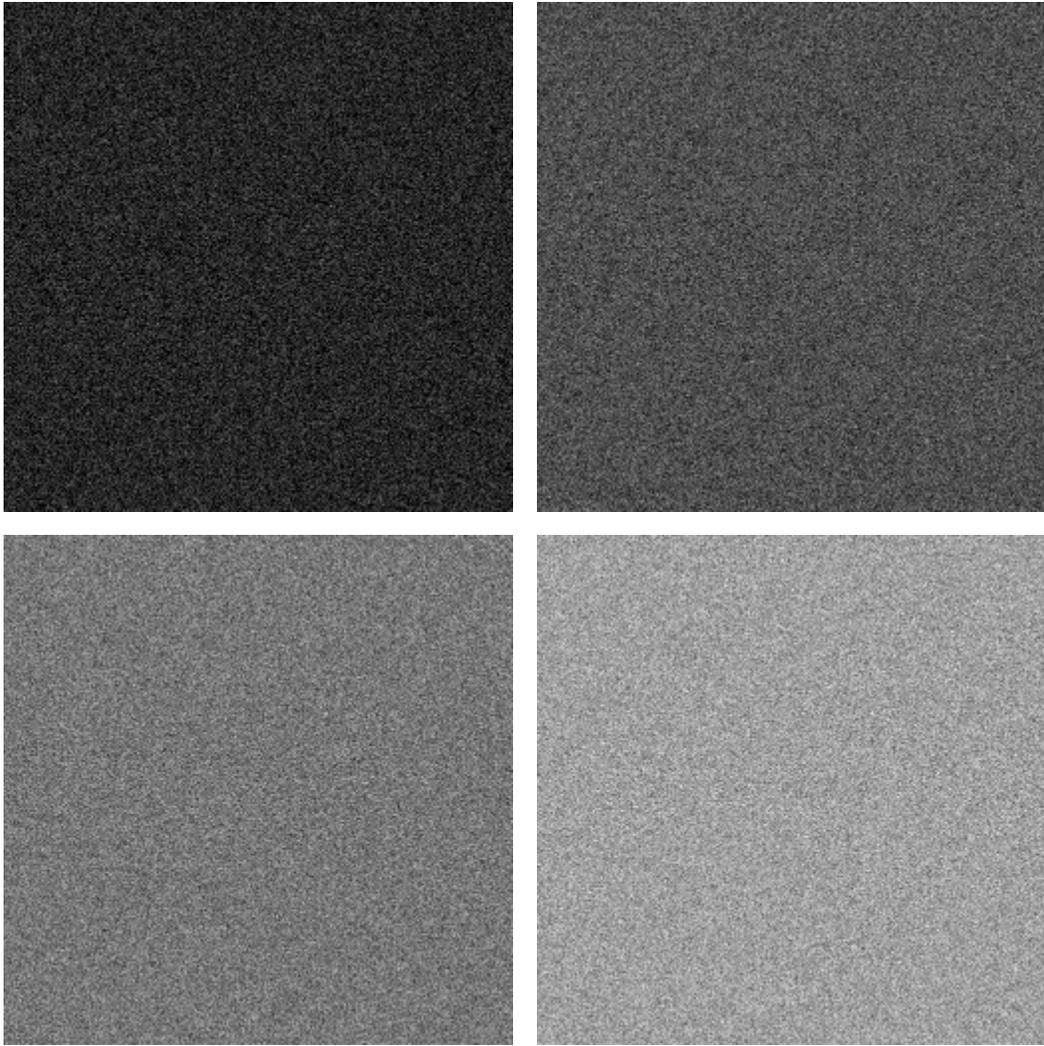
Figura 3.1.2. Detalle ilustrativo de extracción de atributos de la imagen: (a) partiendo de la imagen previamente preprocesada, extraemos ventanas barriendo sobre la imagen al azar, y de cada una de estas ventanas de prueba (b), extraemos sus siete momentos invariantes (c)

3.1.2. Reconocimiento del tipo de letra

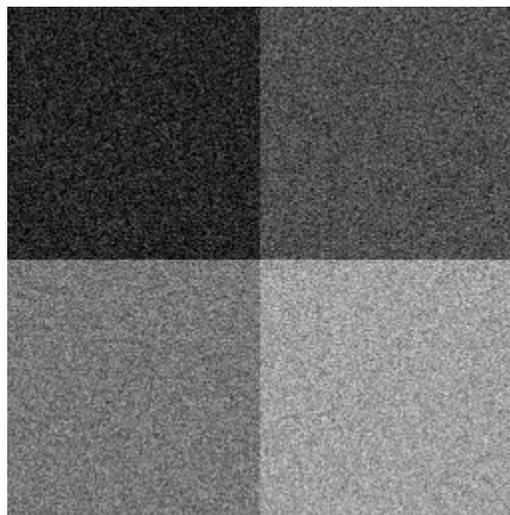
Una vez que tenemos conformada nuestra base de datos, procedemos a aplicar el cálculo de los clasificadores Gaussiano y KNN, pasando a los módulos respectivos nuestra base de datos. Los resultados obtenidos de los experimentos con las clases se muestran en el capítulo siguiente.

Previamente se hicieron experimentos de clasificación con imágenes de ruido aleatorio a 4 diferentes intensidades, y se hizo una imagen de recolección de estas 4 clases así formadas, clases como se muestran en la Figura 3.13, extrayendo una ventana al azar de ésta, y que el clasificador hiciera el reconocimiento.

Una vez validado nuestro clasificador, se procedió con la misma idea a construir una imagen con ocho muestras de las clases de los textos, como se muestra en la Figura 3.14. Esto es siguiendo nuestra cadena de procesamiento, como se describió previamente en la Figura 3.1.



(a)



(b)

Figura 3.13. (a) Clases generadas con ruido al azar y (b) la imagen muestra que se construye a partir de las clases generadas.

Ahora a seguir el punto más importante es determinar el tamaño de la ventana. Como anteriormente se dijo, en un principio tomamos ventanas de prueba de 128 x 128 píxeles, pero el nivel el porcentaje de errores obtenidos era alto, por lo que se decidió tomar el tamaño de 256 x 256 píxeles, para lograr los mejores resultados en la clasificación de la imagen. Para medir la efectividad lo hacemos aplicando los algoritmos de validación, ya descritos, como lo son Resustitución (R), Leave-One-Out (LOO), y Cross-Validation (CV).

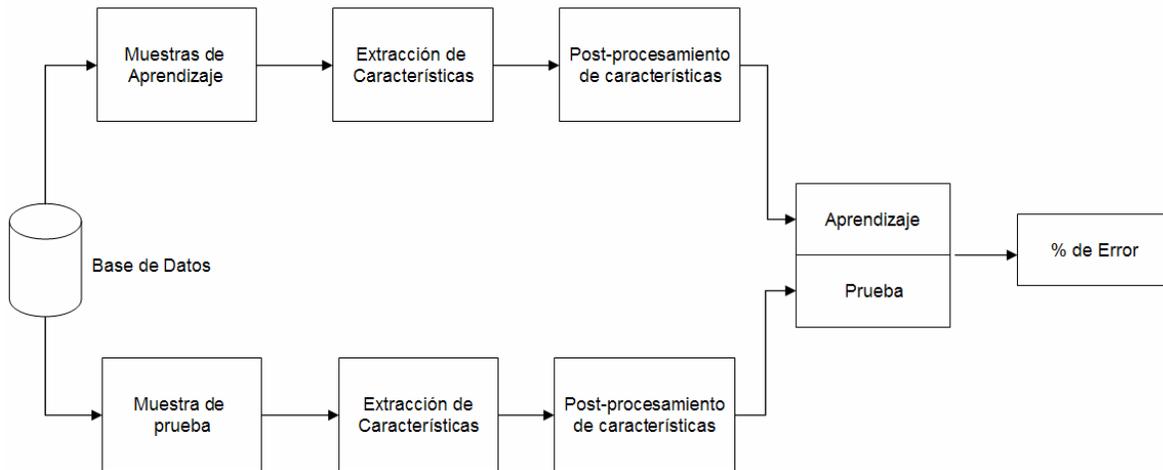


Figura 3.15. Procedimiento de Leave-One-Out.

El procedimiento a seguir en LOO se observa en la Figura 3.15 y consiste en lo siguiente.

Se obtienen las muestras aleatorias de cada clase, formando la base de datos de N muestras. Enseguida las N muestras de cada clase se concatenan en una sola base de datos de tamaño M , que es el n número de clases por N muestras. Así se procede a tomar la primera muestra de la base de datos, ésta se toma como muestra de prueba y las $M-1$ restantes se toman como muestras de aprendizaje. Tanto la muestra de prueba como la de aprendizaje son adquiridas a partir de los atributos de la imagen como se mencionó anteriormente.

Así es que tanto la muestra de prueba como la de aprendizaje se someten a ambos clasificadores (Gaussiano y KNN), y se obtiene en qué clase es clasificada la muestra de prueba. Al terminar de clasificar esta muestra, se toma la segunda muestra de la base de datos M , como muestra de prueba, y nuevamente las $M-1$ restantes se toman como muestras de aprendizaje y se repite el proceso anterior descrito, y así sucesivamente se repite el proceso para las M muestras de la base de datos, hasta haber clasificado el ultimo valor de M .

Los resultados obtenidos se almacenan en una matriz, llamada matriz de confusión, descrita anteriormente, esto con la finalidad de saber cuantas muestras fueron clasificadas correctamente y cuales no. Al tener esta información, se puede medir la eficiencia o el error que se tiene al hacer la clasificación para cierto tamaño de ventana y el número de vecinos (para el algoritmo KNN, que en nuestro caso tomamos $N=3$ vecinos).

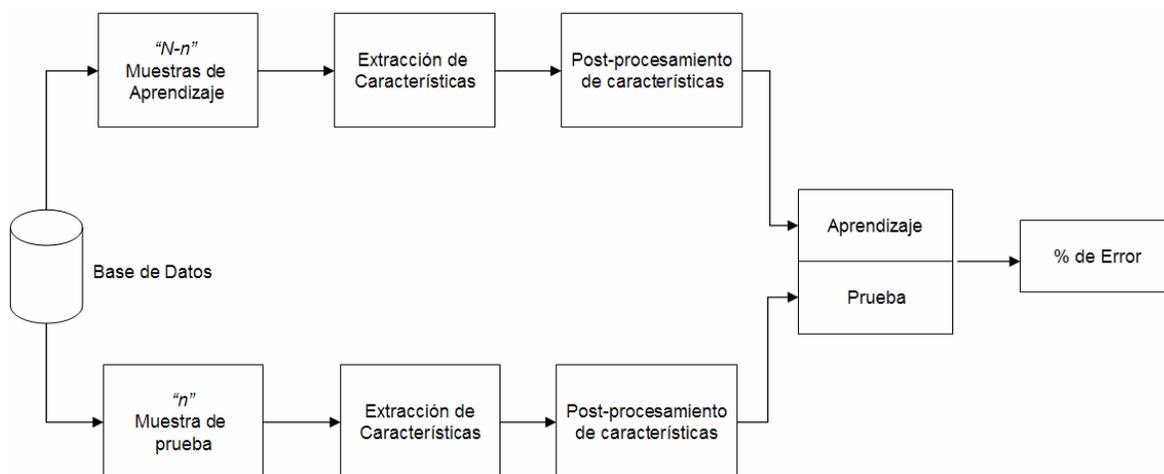


Figura 3.16. Procedimiento de Cross-Validation.

El procedimiento a seguir por el método de validación de CV (Figura 3.16.), se tiene que para llevarlo a cabo, se toma el $n/2$ número de muestras aleatorias de la base de datos de cada clase, estas se toman como muestras de prueba, mientras que las $N-(n/2)$ muestras restantes se toman como muestras de aprendizaje, i.e. de los valores de cada clase tomamos la mitad para clasificar y la otra mitad para aprendizaje. Tanto la muestra de prueba como las de aprendizaje son obtenidas a partir de los atributos de la imagen descritos anteriormente.

Se toman así las muestras de prueba como las de aprendizaje y se someten a los clasificadores (Gaussiano y KNN), y posteriormente obtenemos a través de la matriz de confusión correspondiente cuantas muestras fueron clasificadas correctamente y cuales no.

Este procedimiento se repite las veces que sea necesario, a juicio del usuario, para así obtener un promedio de cuántas fueron bien clasificadas por clase y medir la eficiencia o el error del clasificador.

Hablando finalmente del procedimiento a seguir en el método de Resustitución, es tomar la misma base de datos formada M , a partir de las N muestras, por el n número de clases, como base de aprendizaje y posteriormente como base a clasificar. Como se describió anteriormente, este método de validación es el más idealista, ya que las muestras que clasifica son las mismas de las que aprendió, pero nos proporciona el nivel máximo que podemos esperar de efectividad.

En el capítulo siguiente se muestran los resultados obtenidos a partir de éstos métodos de validación, se presentan por medio de gráficas, y también tomamos en consideración la media de eficiencia de cada clasificador tomando el promedio de la diagonal de la matriz de confusión.

4. Resultados Experimentales.

En esta sección mostramos resultados de experimentos, desde los primeros hechos con los clasificadores, en font regular, y con variaciones de tamaño en puntos (sección 4.1.), luego experimentando con rotaciones de las muestras, y aplicando efectos sobre el font del tipo Cursiva-Itálica, Negrita, y combinación de los efectos en Cursiva con Negrita (sección 4.2 a 4.3). Estos experimentos fueron con un tamaño de ventana de prueba fijo a 256 píxeles, y en la sección 4.4., mostramos los experimentos con las variaciones de tamaño de ventana de prueba, así también mostramos la relación entre tamaño de ventana de prueba y tamaño el font.

4.1. Identificación de tipo de font.

Se hicieron pruebas con los clasificadores, construyendo las matrices de confusión con el método de Resustitución, y con el font regular.

El clasificador *KNN* se tomó con $K=3$.

El cálculo de los momentos invariantes se hizo a variaciones de tomar 7, 6, 5, 4, 3, y 2 de ellos, comenzando con los 2 primeros momentos, y así sucesivamente.

Con el fin de reducir el rango dinámico en el cálculo de los momentos, en algunos experimentos de aplicó el logaritmo de los resultados, esto también se indica.

A continuación se muestran las Matrices de Confusión respectivas considerando una base de aprendizaje con T momentos, tamaño de las ventanas de prueba, y media de probabilidad de cada clasificador.

Resultados para $T = 2$ momentos:

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	65	0	11	0	0	0	0
3	0	0	77	0	1	0	0	54
4	0	34	0	85	0	0	2	0
5	0	0	1	0	84	0	0	23
6	0	0	0	0	0	100	0	0
7	0	1	0	4	0	0	98	0
8	0	0	22	0	15	0	0	23

Figura 4.1. Matriz de confusión con clasificador Gaussiano, con logaritmo.

	1	2	3	4	5	6	7	8
1	98	0	2	0	3	0	0	3
2	0	77	0	16	1	4	11	2
3	1	0	84	1	14	0	4	26
4	0	16	0	81	1	1	8	1
5	1	1	2	0	66	0	0	17
6	0	0	0	0	0	95	0	0
7	0	6	0	2	1	0	77	0
8	0	0	12	0	14	0	0	51

Figura 4.2. Matriz de confusión con clasificador KNN, con logaritmo.

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	81	0	14	0	0	1	0
3	0	0	84	0	5	0	0	25
4	0	18	0	85	0	0	1	0
5	0	0	0	0	87	0	0	14
6	0	0	0	0	0	100	0	0
7	0	1	0	1	0	0	98	0
8	0	0	16	0	8	0	0	61

Figura 4.3. Matriz de confusión con clasificador Gaussiano, sin logaritmo.

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	65	0	14	0	0	0	0
3	0	0	71	0	0	0	0	41
4	0	34	0	84	0	0	4	0
5	0	0	1	0	83	0	0	23
6	0	0	0	0	0	100	0	0
7	0	1	0	2	0	0	96	0
8	0	0	28	0	17	0	0	36

Figura 4.4. Matriz de confusión con clasificador KNN, sin logaritmo.

clasificador / Log	con Log	sin Log
Gaussiano	81.12 %	79.37 %
KNN	79.37 %	87 %

Tabla 4.1. Comparativo de medias de probabilidad con cada clasificador, con y sin logaritmo.

Como podemos observar en las matrices de confusión, se obtiene una mayor probabilidad sin aplicar el logaritmo a los resultados. De aquí el resto de los cálculos los haremos sin logaritmo.

Resultados para T = 3 momentos:

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	70	0	13	0	0	2	0
3	0	0	79	0	0	0	0	33
4	0	29	0	82	0	0	6	0
5	0	0	0	0	86	0	0	18
6	0	0	0	0	0	100	0	0
7	0	1	0	5	0	0	92	0
8	0	0	21	0	14	0	0	49

Figura 4.5. Matriz de confusión con clasificador Gaussiano.

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	82	0	15	0	0	3	0
3	0	0	84	0	2	0	0	18
4	0	18	0	83	0	0	5	0
5	0	0	0	0	87	0	0	10
6	0	0	0	0	0	100	0	0
7	0	0	0	2	0	0	92	0
8	0	0	16	0	11	0	0	72

Figura 4.6. Matriz de confusión con clasificador KNN.

Resultados para T = 4 momentos:

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	64	0	11	0	0	4	0
3	0	0	64	0	1	0	0	34
4	0	32	0	82	0	0	15	0
5	0	0	2	0	86	0	0	21
6	0	0	0	0	0	100	0	0
7	0	4	0	7	0	0	81	0
8	0	0	34	0	13	0	0	45

Figura 4.7. Matriz de confusión con clasificador Gaussiano.

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	78	0	16	0	0	10	0
3	0	0	79	0	5	0	0	25
4	0	22	0	80	0	0	6	0
5	0	0	2	0	93	0	0	13
6	0	0	0	0	0	100	0	0
7	0	0	0	4	0	0	84	0
8	0	0	19	0	2	0	0	62

Figura 4.8. Matriz de confusión con clasificador KNN.

Resultados para T = 5 momentos:

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	80	0	10	0	0	3	0
3	0	0	84	0	3	0	0	24
4	0	20	0	87	0	0	4	0
5	0	0	0	0	89	0	0	19
6	0	0	0	0	0	100	0	0
7	0	0	0	3	0	0	93	0
8	0	0	16	0	8	0	0	57

Figura 4.9. Matriz de confusión con clasificador Gaussiano.

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	62	0	7	0	0	4	0
3	0	0	78	0	1	0	0	30
4	0	37	0	85	0	0	4	0
5	0	0	3	0	89	0	0	42
6	0	0	0	0	0	100	0	0
7	0	1	0	8	0	0	92	0
8	0	0	19	0	10	0	0	28

Figura 4.10. Matriz de confusión con clasificador KNN.

Resultados para T = 6 momentos:

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	57	0	7	0	0	4	0
3	0	0	84	0	3	0	0	45
4	0	42	0	86	0	0	10	0
5	0	0	9	0	92	0	0	29
6	0	0	0	0	0	100	0	0
7	0	1	0	7	0	0	86	1
8	0	0	7	0	5	0	0	25

Figura 4.11. Matriz de confusión con clasificador Gaussiano.

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	72	0	8	0	0	1	0
3	0	0	82	0	6	0	0	28
4	0	27	0	87	0	0	4	0
5	0	0	4	0	88	0	0	13
6	0	0	0	0	0	100	0	0
7	0	1	0	5	0	0	95	0
8	0	0	14	0	6	0	0	59

Figura 4.12. Matriz de confusión con clasificador KNN.

Resultados para T = 7 momentos:

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	83	0	15	0	0	1	0
3	0	0	82	0	2	0	0	17
4	0	16	0	83	0	0	4	0
5	0	0	4	0	87	0	0	16
6	0	0	0	0	0	100	0	0
7	0	1	0	2	0	0	95	0
8	0	0	14	0	11	0	0	67

Figura 4.13. Matriz de confusión con clasificador Gaussiano.

	1	2	3	4	5	6	7	8
1	100	0	0	0	0	0	0	0
2	0	51	1	9	0	0	6	0
3	0	0	54	0	2	0	0	25
4	0	47	0	90	0	0	8	0
5	0	0	11	0	83	0	0	32
6	0	0	0	0	0	100	0	0
7	0	2	0	1	0	0	86	0
8	0	0	34	0	15	0	0	43

Figura 4.14. Matriz de confusión con clasificador KNN.

	2	3	4	5	6	7
Gaussiano	79.37	82.25	77.75	79.25	78.75	75.87
KNN	87.00	87.50	84.50	86.25	85.37	87.12

Tabla 4.2. Comparativo de medias de probabilidad con cada clasificador, y con los momentos considerados.

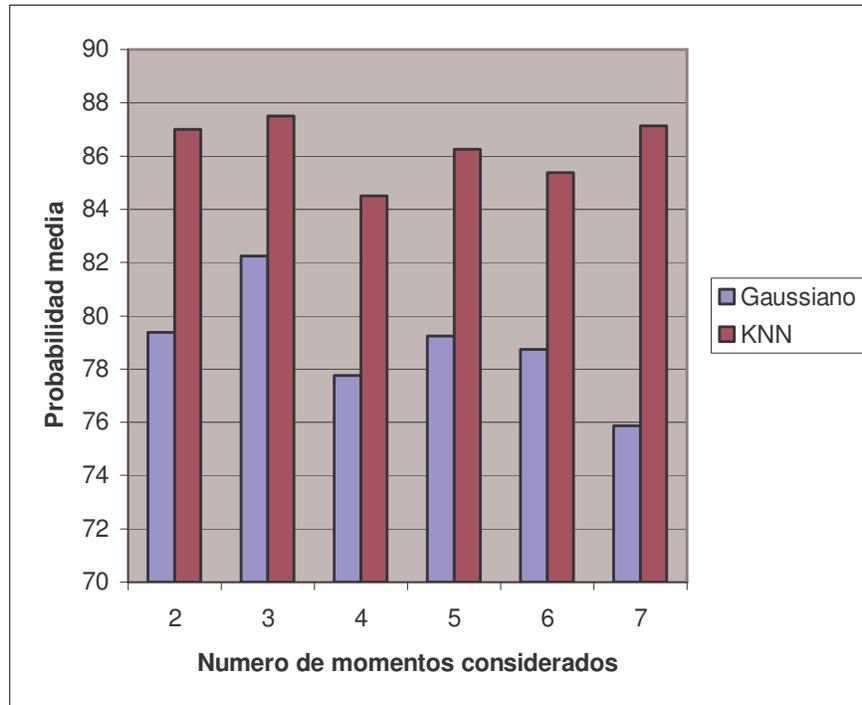


Figura 4.15. Grafica de barras de probabilidades por clasificador y por número de momentos considerados.

Como conclusión de estos experimentos, podemos ver que considerando solo 3 momentos alcanzamos la más alta probabilidad para ambos clasificadores, bajo el método de Resustitución.

4.2. Identificación de tipo de font con rotaciones de: 0°, 45°, 90°, 135°, a 6, 8 10, 12 puntos, para el caso de letra regular.

Los siguientes experimentos se hacen sobre los siguientes casos de tipos de font, tamaño en puntos y rotaciones, como se muestra en la Tabla 4.3. En estos casos el tipo de letra es regular, sin efecto alguno (no-cursiva, no-negrita).

N°	Font	Escalas				Rotación			
		12	10	8	6	0°	45°	90°	135°
1	Courier New	✓	✓	✓	✓	✓	✓	✓	✓
2	Arial	✓	✓	✓	✓	✓	✓	✓	✓
3	Bookman Old Style	✓	✓	✓	✓	✓	✓	✓	✓
4	Franklin Ghotic Médium	✓	✓	✓	✓	✓	✓	✓	✓
5	Comic Sans MS	✓	✓	✓	✓	✓	✓	✓	✓
6	Impact	✓	✓	✓	✓	✓	✓	✓	✓
7	Modern	✓	✓	✓	✓	✓	✓	✓	✓
8	Times New Roman	✓	✓	✓	✓	✓	✓	✓	✓

Tabla 4.3. Todos los casos considerados en las pruebas, por cada escala se toman las 4 medidas de rotación (e.g. 12 pts. En 0°, 45°, 90° y 135°).

Para cada uno de estos casos de se hacen las pruebas de los clasificadores Gaussiano y KNN ($K=3$).

En esta serie de cálculos se toman en cuenta los 7 momentos *invariantes de Hue*, con las muestras sin ruido, de font normal, con rotaciones a 0°, 45°, 90° y 135°, sin distorsión, como se muestran en la Figura 4.16 unos ejemplos de las clases.

4.2.1. Caso: 12 pts, rotación 0°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	98	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	95	0	0	9	0	0	0		0	98	0	0	11	0	0	0
BookmanOS	2	0	36	0	2	0	6	8		0	0	80	0	0	0	21	24
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	5	3	0	89	0	0	0		0	2	0	0	89	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	21	0	0	0	73	48		0	0	4	0	0	0	68	16
Times N. R.	0	0	40	0	0	0	21	44		0	0	16	0	0	0	11	60

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	47	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	43	0	0	7	0	0	0		0	47	0	0	6	0	0	0
BookmanOS	3	0	15	0	1	0	1	7		0	0	39	0	0	0	7	11
FranklinGM	0	1	0	50	1	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	6	2	0	41	0	0	0		0	3	0	0	44	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	9	0	0	0	29	11		0	0	1	0	0	0	36	5
Times N. R.	0	0	24	0	0	0	20	32		0	0	10	0	0	0	7	34

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	97	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	91	0	0	11	0	0	0		0	93	0	0	16	0	0	0
BookmanOS	3	0	34	0	2	0	8	9		0	0	61	0	0	0	28	41
FranklinGM	0	1	0	100	1	0	0	0		0	0	0	99	0	0	0	2
Comic Sans	0	8	4	0	86	0	0	0		0	7	0	1	84	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	22	0	0	0	67	50		0	0	10	0	0	0	41	24
Times N. R.	0	0	40	0	0	0	25	41		0	0	29	0	0	0	31	33

Gauss

KNN

La Tabla 4.4., muestra el resumen de las probabilidades medias por clasificador por método de validación.

	Gauss	KNN (N=3)
Resustitución	79.37	86.87
Cross-Validation	76.75	87.50
L.O.O.	77.00	76.37

Tabla 4.4. Resumen de probabilidades medias para el caso de 12 puntos, 0° rotación.

4.2.2. Caso: 10 pts, rotación 0°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	98	0	1	0	0	0	0		0	98	0	3	0	0	0	0
BookmanOS	0	0	88	0	0	0	77	24		0	0	72	0	0	0	27	10
FranklinGM	0	2	0	99	0	0	0	0		0	2	0	97	0	0	0	0
Comic Sans	0	0	0	0	100	0	3	5		0	0	0	0	99	0	0	1
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	4	0	0	0	17	4		0	0	20	0	0	0	69	1
Times N. R.	0	0	8	0	0	0	3	67		0	0	8	0	1	0	4	88

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	48	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	49	0	2	2	0	0	0		0	50	0	1	0	0	0	0
BookmanOS	2	0	35	0	0	0	35	10		0	0	37	0	0	0	14	5
FranklinGM	0	1	0	48	0	0	0	0		0	0	0	49	0	0	0	0
Comic Sans	0	0	1	0	48	0	2	4		0	0	0	0	50	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	7	0	0	0	10	1		0	0	10	0	0	0	34	0
Times N. R.	0	0	7	0	0	0	3	35		0	0	3	0	0	0	2	45

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	98	0	0	0	0	0	0	0		10	0	0	0	0	0	0	0
Arial	0	96	0	3	2	0	0	0		0	97	0	3	0	0	0	0
BookmanOS	2	0	81	0	0	0	79	25		0	0	54	0	0	0	50	17
FranklinGM	0	4	0	97	0	0	0	0		0	3	1	97	0	0	0	0
Comic Sans	0	0	0	0	98	0	4	6		0	0	0	0	97	0	0	2
Impact	0	0	0	0	0	100	0	0		0	0	0	0	1	98	0	0
Modern	0	0	8	0	0	0	14	4		0	0	31	0	0	2	40	1
Times N. R.	0	0	11	0	0	0	3	65		0	0	14	0	2	0	10	80

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	83.62	90.37
Cross-Validation	80.75	91.25
L.O.O.	81.125	82.87

Tabla 4.5. Resumen de probabilidades medias para el caso de 10 puntos, 0° rotación.

4.2.3. Caso: 8 pts, rotación 0°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	99	0	0	0	0	0	0		0	100	0	0	0	0	0	0
BookmanOS	0	0	99	0	0	0	0	0		0	0	99	0	0	0	0	1
FranklinGM	0	1	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	0	0	0	100	0	1	0		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	81	14		0	0	0	0	0	0	92	15
Times N. R.	0	0	1	0	0	0	18	86		0	0	1	0	0	0	8	84

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	46	0	0	0	0	0	0		0	50	0	0	0	0	0	0
BookmanOS	0	0	37	0	0	0	0	1		0	0	50	0	0	0	0	0
FranklinGM	0	4	0	50	0	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	0	2	0	50	0	4	3		0	0	0	0	50	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	0	0	0	0	35	9		0	0	0	0	0	0	44	8
Times N. R.	0	0	11	0	0	0	11	37		0	0	0	0	0	0	6	42

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	99	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	95	0	0	3	0	0	0		0	100	0	0	1	0	0	0
BookmanOS	1	0	94	0	0	0	0	0		0	0	97	0	0	0	0	3
FranklinGM	0	5	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	0	2	0	97	0	3	1		0	0	0	0	98	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	1	100	0	0
Modern	0	0	1	0	0	0	76	17		0	0	0	0	0	0	73	21
Times N. R.	0	0	3	0	0	0	21	82		0	0	3	0	0	0	27	76

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	95.62	96.87
Cross-Validation	88.75	96.50
L.O.O.	92.87	93.00

Tabla 4.6. Resumen de probabilidades medias para el caso de 8 puntos, 0° rotación.

4.2.4. Caso: 6 pts, rotación 0°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	94	0	0	13	0	0	0		0	97	0	0	14	0	0	0
BookmanOS	0	0	100	0	0	0	0	0		0	0	100	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	6	0	0	87	0	0	0		0	3	0	0	86	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	87	24		0	0	0	0	0	0	90	17
Times N. R.	0	0	0	0	0	0	13	76		0	0	0	0	0	0	10	83

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	47	0	0	2	0	0	0		0	49	0	0	4	0	0	0
BookmanOS	0	0	47	0	0	0	0	1		0	0	50	0	0	0	0	0
FranklinGM	0	0	0	50	1	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	3	0	0	47	0	0	0		0	1	0	0	46	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	1	0	0	0	42	22		0	0	0	0	0	0	45	9
Times N. R.	0	0	2	0	0	0	8	27		0	0	0	0	0	0	5	41

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	94	0	0	15	0	0	0		0	93	0	0	13	0	0	0
BookmanOS	0	0	100	0	0	0	0	0		0	0	99	0	0	0	0	2
FranklinGM	0	0	0	98	0	0	0	0		0	0	0	100	1	0	0	0
Comic Sans	0	6	0	2	85	0	0	0		0	7	0	0	86	0	0	1
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	72	27		0	0	1	0	0	0	85	25
Times N. R.	0	0	0	0	0	0	28	73		0	0	0	0	0	0	15	72

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	93.00	94.50
Cross-Validation	90.00	95.25
L.O.O.	91.87	90.25

Tabla 4.7. Resumen de probabilidades medias para el caso de 6 puntos, 0° rotación.

4.2.5. Caso: 12 pts, rotación 45°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	96	0	0	3	0	0	0		0	97	0	0	3	0	0	0
BookmanOS	0	0	48	0	0	0	7	7		0	0	81	0	0	0	17	17
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	4	1	0	97	0	0	0		0	3	0	0	97	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	5	0	0	0	58	12		0	0	3	0	0	0	70	8
Times N. R.	0	0	46	0	0	0	35	81		0	0	16	0	0	0	13	75

Gauss KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	49	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	47	0	0	1	0	0	0		0	49	0	0	2	0	0	0
BookmanOS	0	0	21	0	0	0	5	14		0	0	42	0	0	0	9	7
FranklinGM	0	0	0	50	1	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	3	10	0	48	0	0	1		0	1	0	0	48	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	1	0	9	0	0	0	29	8		0	0	2	0	0	0	36	7
Times N. R.	0	0	10	0	0	0	16	27		0	0	6	0	0	0	5	36

Gauss KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	98	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	92	0	0	4	0	0	0		0	96	0	0	6	0	0	0
BookmanOS	0	0	43	0	0	0	10	11		0	0	67	0	0	0	15	45
FranklinGM	0	0	0	100	1	0	0	0		0	0	0	99	0	0	0	0
Comic Sans	0	8	4	0	95	0	0	1		0	4	0	1	92	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	2	100	0	0
Modern	2	0	6	0	0	0	53	13		0	0	7	0	0	0	60	13
Times N. R.	0	0	47	0	0	0	37	75		0	0	26	0	0	0	25	42

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	85.00	90.00
Cross-Validation	80.25	90.25
L.O.O.	82.00	82.00

Tabla 4.8. Resumen de probabilidades medias para el caso de 12 puntos, 45° rotación.

4.2.6. Caso: 10 pts, rotación 45°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	100	0	1	1	0	0	0		0	100	0	0	0	0	0	0
BookmanOS	0	0	27	0	0	0	9	7		0	0	54	0	0	0	14	12
FranklinGM	0	0	0	99	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	0	1	0	98	0	1	1		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	45	0	0	0	86	8		0	0	27	0	0	0	83	1
Times N. R.	0	0	27	0	1	0	4	84		0	0	19	0	0	0	3	87

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	48	0	0	1	0	0	0		0	50	0	0	0	0	0	0
BookmanOS	0	0	17	0	0	0	14	5		0	0	30	0	0	0	7	5
FranklinGM	0	2	0	50	0	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	0	6	0	48	0	2	4		0	0	0	0	50	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	17	0	0	0	32	8		0	0	14	0	0	0	40	0
Times N. R.	0	0	10	0	1	0	2	33		0	0	6	0	0	0	3	45

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	99	0	1	2	0	0	0		0	98	0	0	0	0	0	0
BookmanOS	0	0	23	0	0	0	14	12		0	2	36	0	0	0	43	26
FranklinGM	0	1	0	99	0	0	0	0		0	0	0	100	0	0	0	2
Comic Sans	0	0	3	0	97	0	1	1		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	46	0	0	0	81	8		0	0	39	0	0	0	53	1
Times N. R.	0	0	28	0	1	0	4	79		0	0	25	0	0	0	4	71

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	86.75	90.50
Cross-Validation	82.00	91.25
L.O.O.	84.75	82.25

Tabla 4.9. Resumen de probabilidades medias para el caso de 10 puntos, 45° rotación.

4.2.7. Caso: 8 pts, rotación 45°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	100	0	0	0	0	0	0		0	100	0	0	0	0	0	0
BookmanOS	0	0	100	0	0	0	0	0		0	0	100	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	0	0	0	100	0	2	1		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	94	8		0	0	0	0	0	0	100	9
Times N. R.	0	0	0	0	0	0	4	91		0	0	0	0	0	0	0	91

Gauss KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	48	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	44	0	0	0	0	0	0		0	50	0	0	0	0	0	0
BookmanOS	0	0	48	0	0	0	0	1		0	0	50	0	0	0	0	0
FranklinGM	0	2	0	50	3	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	2	4	0	0	47	0	3	3		0	0	0	0	50	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	0	0	0	0	41	3		0	0	0	0	0	0	48	3
Times N. R.	0	0	2	0	0	0	6	43		0	0	0	0	0	0	2	47

Gauss KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	99	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	97	0	0	0	0	0	0		0	100	0	0	0	0	0	0
BookmanOS	0	0	96	0	0	0	0	0		0	0	100	0	0	0	0	0
FranklinGM	0	1	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	1	2	1	0	100	0	3	3		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	98	0	0
Modern	0	0	0	0	0	0	92	12		0	0	0	0	0	2	96	11
Times N. R.	0	0	3	0	0	0	5	85		0	0	0	0	0	0	4	89

Gauss KNN

	Gauss	KNN (N=3)
Resustitución	98.12	98.87
Cross-Validation	92.75	98.75
L.O.O.	96.125	97.875

Tabla 4.10. Resumen de probabilidades medias para el caso de 8 puntos, 45° rotación.

4.2.8. Caso: 6 pts, rotación 45°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
Arial	0	92	0	0	5	0	0	0	0	98	0	0	2	0	0	0	0
BookmanOS	0	0	100	0	0	0	0	0	0	0	100	0	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0	0	0	0	100	0	0	0	0	0
Comic Sans	0	8	0	0	95	0	0	0	0	2	0	0	98	0	0	0	0
Impact	0	0	0	0	0	100	0	0	0	0	0	0	0	100	0	0	0
Modern	0	0	0	0	0	0	77	7	0	0	0	0	0	0	91	3	0
Times N. R.	0	0	0	0	0	0	23	93	0	0	0	0	0	0	9	97	0

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0
Arial	0	49	0	0	7	0	0	0	0	49	0	0	2	0	0	0	0
BookmanOS	0	0	49	0	0	0	0	1	0	0	50	0	0	0	0	0	0
FranklinGM	0	0	0	50	0	0	0	0	0	0	0	50	0	0	0	0	0
Comic Sans	0	1	0	0	43	0	0	0	0	1	0	0	48	0	0	0	0
Impact	0	0	0	0	0	50	0	0	0	0	0	0	0	50	0	0	0
Modern	0	0	1	0	0	0	41	7	0	0	0	0	0	0	44	3	0
Times N. R.	0	0	0	0	0	0	9	42	0	0	0	0	0	0	6	47	0

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	91	0	0	8	0	0	0		0	98	0	0	3	0	0	0
BookmanOS	0	0	99	0	0	0	0	0		0	0	98	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	2	100	0	0	0	0
Comic Sans	0	9	0	0	92	0	0	0		0	2	0	0	97	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	1	0	0	0	77	13		0	0	0	0	0	0	85	9
Times N. R.	0	0	0	0	0	0	23	87		0	0	0	0	0	0	15	91

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	94.625	98
Cross-Validation	93.5	97
L.O.O.	93.25	96.125

Tabla 4.11. Resumen de probabilidades medias para el caso de 6 puntos, 45° rotación.

4.2.9. Caso: 12 pts, rotación 90°.

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	99	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	97	0	0	7	0	0	0		0	99	0	0	4	0	0	0
BookmanOS	1	0	39	0	0	0	9	6		0	0	85	0	0	0	16	27
FranklinGM	0	1	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	2	1	0	93	0	0	0		0	1	0	0	96	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	8	0	0	0	43	9		0	0	8	0	0	0	70	10
Times N. R.	0	0	52	0	0	0	48	85		0	0	7	0	0	0	14	63

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	49	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	39	0	0	8	0	0	0		0	49	0	0	1	0	0	0
BookmanOS	1	0	19	0	0	0	3	1		0	0	41	0	0	0	9	14
FranklinGM	0	9	0	49	3	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	2	2	0	39	0	0	0		0	1	0	0	49	0	0	0
Impact	0	0	0	1	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	4	0	0	0	18	5		0	0	5	0	0	0	38	3
Times N. R.	0	0	25	0	0	0	29	44		0	0	4	0	0	0	3	33

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	97	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	95	0	0	7	0	0	0		0	99	0	0	5	0	0	0
BookmanOS	3	0	36	0	0	0	11	8		0	0	66	0	0	0	29	35
FranklinGM	0	2	0	100	0	0	0	0		0	0	0	100	0	0	2	1
Comic Sans	0	3	3	0	93	0	0	0		0	1	0	0	95	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	99	0	0
Modern	0	0	9	0	0	0	38	15		0	0	13	0	0	1	44	24
Times N. R.	0	0	52	0	0	0	51	77		0	0	21	0	0	0	25	40

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	82	89.125
Cross-Validation	76.75	90
L.O.O.	79.5	80.375

Tabla 4.12. Resumen de probabilidades medias para el caso de 12 puntos, 90° rotación.

4.2.10. Caso: 10 pts, rotación 90°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	94	0	1	1	0	0	0		0	97	0	0	0	0	0	0
BookmanOS	0	0	57	0	0	0	12	9		0	0	67	0	0	0	19	14
FranklinGM	0	6	0	99	0	0	0	0		0	3	0	100	0	0	0	0
Comic Sans	0	0	1	0	99	0	1	2		0	0	0	0	100	0	0	1
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	28	0	0	0	85	6		0	0	22	0	0	0	79	1
Times N. R.	0	0	14	0	0	0	2	83		0	0	11	0	0	0	2	84

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	45	0	1	1	0	0	0		0	49	0	0	0	0	0	0
BookmanOS	0	0	30	0	0	0	6	7		0	0	40	0	0	0	9	5
FranklinGM	0	4	0	49	0	0	0	0		0	1	0	50	0	0	0	0
Comic Sans	0	0	2	0	49	0	0	3		0	0	0	0	50	0	0	0
Impact	0	1	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	13	0	0	0	43	6		0	0	8	0	0	0	40	0
Times N. R.	0	0	5	0	0	0	1	34		0	0	2	0	0	0	1	45

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	99	0	0	0	0	1	0	0		100	0	0	0	0	0	0	0
Arial	0	92	0	5	2	0	0	0		0	97	0	5	0	0	0	0
BookmanOS	1	0	54	0	0	0	14	14		0	1	48	0	0	0	36	27
FranklinGM	0	8	0	95	0	0	0	0		0	2	0	95	0	0	0	0
Comic Sans	0	0	3	0	98	0	1	3		0	0	0	0	100	0	0	1
Impact	0	0	0	0	0	99	0	0		0	0	0	0	0	99	0	0
Modern	0	0	29	0	0	0	83	6		0	0	33	0	0	1	58	1
Times N. R.	0	0	14	0	0	0	2	77		0	0	19	0	0	0	6	71

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	89.625	90.875
Cross-Validation	87.5	93.5
L.O.O.	87.125	83.5

Tabla 4.13. Resumen de probabilidades medias para el caso de 10 puntos, 90° rotación.

4.2.11. Caso: 8 pts, rotación 90°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
Arial	0	100	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
BookmanOS	0	0	96	0	0	0	0	3	0	0	97	0	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0	0	0	0	100	0	0	0	0	0
Comic Sans	0	0	0	0	100	0	1	1	0	0	0	0	100	0	0	0	0
Impact	0	0	0	0	0	100	0	0	0	0	0	0	0	100	0	0	0
Modern	0	0	0	0	0	0	91	20	0	0	0	0	0	0	96	13	
Times N. R.	0	0	4	0	0	0	8	76	0	0	3	0	0	0	4	87	

Gauss KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	48	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0
Arial	0	48	0	0	0	0	0	0	0	49	0	0	0	0	0	0	0
BookmanOS	1	0	46	0	0	0	0	3	0	0	50	0	0	0	0	0	3
FranklinGM	0	2	0	50	0	0	0	0	0	0	0	50	0	0	0	0	0
Comic Sans	1	0	0	0	50	0	0	0	0	1	0	0	50	0	0	0	0
Impact	0	0	0	0	0	50	0	0	0	0	0	0	0	50	0	0	0
Modern	0	0	0	0	0	0	42	9	0	0	0	0	0	0	46	6	
Times N. R.	0	0	4	0	0	0	8	38	0	0	0	0	0	0	4	41	

Gauss KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	98	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	96	0	0	1	0	0	0		0	99	0	0	0	0	0	0
BookmanOS	2	0	93	0	0	0	0	4		0	0	95	0	0	0	0	4
FranklinGM	0	4	0	100	0	0	0	0		0	0	1	100	0	0	0	0
Comic Sans	0	0	0	0	99	0	3	2		0	1	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	98	0	0
Modern	0	0	0	0	0	0	89	23		0	0	0	0	0	2	90	17
Times N. R.	0	0	7	0	0	0	8	71		0	0	4	0	0	0	10	79

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	95.375	97.5
Cross-Validation	93	96.5
L.O.O.	93.25	95.125

Tabla 4.14. Resumen de probabilidades medias para el caso de 8 puntos, 90° rotación.

4.2.12. Caso: 6 pts, rotación 90°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	94	0	0	25	0	0	0		0	93	0	0	12	0	0	0
BookmanOS	0	0	99	0	0	0	0	0		0	0	100	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	6	0	0	75	0	0	0		0	7	0	0	88	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	55	7		0	0	0	0	0	0	80	10
Times N. R.	0	0	1	0	0	0	45	93		0	0	0	0	0	0	20	90

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	47	0	0	13	0	0	0		0	49	0	0	5	0	0	0
BookmanOS	0	0	47	0	0	0	0	0		0	0	50	0	0	0	0	0
FranklinGM	0	0	0	50	1	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	3	0	0	36	0	1	0		0	1	0	0	45	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	1	0	0	0	31	9		0	0	0	0	0	0	42	3
Times N. R.	0	0	2	0	0	0	18	41		0	0	0	0	0	0	8	47

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	90	0	0	25	0	0	0		0	92	0	0	14	0	0	0
BookmanOS	0	0	98	0	0	0	0	1		0	0	100	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	10	1	0	75	0	1	0		0	8	0	0	86	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	53	13		0	0	0	0	0	0	66	29
Times N. R.	0	0	1	0	0	0	46	86		0	0	0	0	0	0	34	71

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	89.5	93.875
Cross-Validation	88	95.75
L.O.O.	87.75	89.375

Tabla 4.15. Resumen de probabilidades medias para el caso de 6 puntos, 90° rotación.

4.2.13. Caso: 12 pts, rotación 135°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	94	0	0	6	0	0	0		0	97	0	0	1	0	0	0
BookmanOS	0	0	57	0	0	0	8	10		0	0	83	0	0	0	19	20
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	6	0	0	94	0	0	2		0	3	0	0	99	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	10	0	0	0	50	10		0	0	5	0	0	0	69	13
Times N. R.	0	0	33	0	0	0	42	78		0	0	12	0	0	0	12	67

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	46	0	0	3	0	0	0		0	49	0	0	1	0	0	0
BookmanOS	0	0	25	0	0	0	3	6		0	0	46	0	0	0	11	14
FranklinGM	0	1	0	49	1	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	3	2	0	46	0	1	2		0	1	0	0	49	0	0	0
Impact	0	0	0	1	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	8	0	0	0	15	5		0	0	1	0	0	0	35	7
Times N. R.	0	0	15	0	0	0	31	37		0	0	3	0	0	0	4	29

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	92	0	0	6	0	0	0		0	97	0	0	4	0	0	0
BookmanOS	0	0	53	0	1	0	11	12		0	0	67	0	0	0	24	43
FranklinGM	0	0	0	99	0	0	0	0		0	0	1	100	0	0	0	0
Comic Sans	0	8	2	0	93	0	2	2		0	3	0	0	94	0	0	0
Impact	0	0	0	1	0	100	0	0		0	0	0	0	2	100	0	0
Modern	0	0	11	0	0	0	44	11		0	0	11	0	0	0	53	20
Times N. R.	0	0	34	0	0	0	43	75		0	0	21	0	0	0	23	37

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	84.125	89.375
Cross-Validation	79.5	89.5
L.O.O.	82	81

Tabla 4.16. Resumen de probabilidades medias para el caso de 12 puntos, 135° rotación.

4.2.14. Caso: 10 pts, rotación 135°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	100	0	0	0	0	0	0		0	100	0	0	0	0	0	0
BookmanOS	0	0	23	0	0	0	10	7		0	0	68	0	0	0	22	13
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	0	0	0	99	0	1	0		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	45	0	0	0	66	1		0	0	23	0	0	0	73	0
Times N. R.	0	0	32	0	1	0	23	92		0	0	9	0	0	0	5	87

Gauss KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	48	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	50	0	2	1	0	0	0		0	50	0	0	0	0	0	0
BookmanOS	2	0	31	0	0	0	35	5		0	0	34	0	0	0	9	3
FranklinGM	0	0	0	48	0	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	0	3	0	49	0	1	0		0	0	0	0	50	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	9	0	0	0	10	5		0	0	12	0	0	0	38	2
Times N. R.	0	0	7	0	0	0	4	40		0	0	4	0	0	0	3	45

Gauss KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	97	0	1	5	0	0	0		0	99	0	1	0	0	0	0
BookmanOS	0	0	18	0	0	0	13	11		0	0	46	0	0	0	46	20
FranklinGM	0	3	0	99	0	0	0	0		0	1	0	99	0	0	2	0
Comic Sans	0	0	2	0	94	0	1	0		0	0	0	0	98	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	2	100	0	0
Modern	0	0	48	0	0	0	62	2		0	0	41	0	0	0	43	5
Times N. R.	0	0	32	0	1	0	24	87		0	0	13	0	0	0	9	75

Gauss KNN

	Gauss	KNN (N=3)
Resustitución	85.00	91.00
Cross-Validation	81.50	91.75
L.O.O.	82.12	82.50

Tabla 4.17. Resumen de probabilidades medias para el caso de 10 puntos, 135° rotación.

4.2.15. Caso: 8 pts, rotación 135°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	99	0	0	0	0	0	0		0	100	0	0	0	0	0	0
BookmanOS	0	0	100	0	0	0	0	3		0	0	99	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	1	0	0	100	0	1	0		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	94	15		0	0	0	0	0	0	98	3
Times N. R.	0	0	0	0	0	0	5	82		0	0	1	0	0	0	2	97

Gauss KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	47	0	0	0	0	0	0		0	50	0	0	0	0	0	0
BookmanOS	0	0	49	0	0	0	0	2		0	0	50	0	0	0	0	0
FranklinGM	0	1	0	50	0	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	2	0	0	50	0	1	3		0	0	0	0	50	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	0	0	0	0	45	5		0	0	0	0	0	0	48	2
Times N. R.	0	0	1	0	0	0	4	40		0	0	0	0	0	0	2	48

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin GM	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	97	0	0	0	0	0	0		0	100	0	0	0	0	0	0
BookmanOS	0	0	98	0	0	0	0	3		0	0	99	0	0	0	0	0
FranklinGM	0	2	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	1	0	0	100	0	4	3		0	0	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	98	0	0
Modern	0	0	0	0	0	0	87	15		0	0	0	0	0	2	91	10
Times N. R.	0	0	2	0	0	0	9	79		0	0	1	0	0	0	9	90

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	96.87	99.25
Cross-Validation	95.25	99.00
L.O.O.	95.12	97.25

Tabla 4.18. Resumen de probabilidades medias para el caso de 8 puntos, 135° rotación.

4.2.16. Caso: 6 pts, rotación 135°

Matrices de confusión por el método de Resustitución:

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	93	0	0	1	0	0	0		0	98	0	0	0	0	0	0
BookmanOS	0	0	100	0	0	0	0	0		0	0	100	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	7	0	0	99	0	0	0		0	2	0	0	100	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	0	100	0	0
Modern	0	0	0	0	0	0	72	11		0	0	0	0	0	0	89	5
Times N. R.	0	0	0	0	0	0	28	89		0	0	0	0	0	0	11	95

Gauss

KNN

Matrices de confusión por el método de Cross-Validation (CV):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	50	0	0	0	0	0	0	0		50	0	0	0	0	0	0	0
Arial	0	46	0	0	0	0	0	0		0	49	0	0	1	0	0	0
BookmanOS	0	0	50	0	0	0	0	0		0	0	50	0	0	0	0	0
FranklinGM	0	0	0	50	0	0	0	0		0	0	0	50	0	0	0	0
Comic Sans	0	4	0	0	50	0	1	0		0	1	0	0	49	0	0	0
Impact	0	0	0	0	0	50	0	0		0	0	0	0	0	50	0	0
Modern	0	0	0	0	0	0	41	11		0	0	0	0	0	0	44	3
Times N. R.	0	0	0	0	0	0	8	39		0	0	0	0	0	0	6	47

Gauss

KNN

Matrices de confusión por el método de Leave-One-Out (LOO):

	Courier New	Arial	Bookman OS	Franklin G M	Comic Sans	Impact	Modern	Times N R		Courier New	Arial	Bookman O S	Franklin GM	Comic Sans	Impact	Modern	Times N R
Courier New	100	0	0	0	0	0	0	0		100	0	0	0	0	0	0	0
Arial	0	93	0	0	5	0	0	0		0	96	0	0	2	0	0	0
BookmanOS	0	0	100	0	0	0	0	0		0	0	100	0	0	0	0	0
FranklinGM	0	0	0	100	0	0	0	0		0	0	0	100	0	0	0	0
Comic Sans	0	7	0	0	95	0	1	0		0	4	0	0	97	0	0	0
Impact	0	0	0	0	0	100	0	0		0	0	0	0	1	100	0	0
Modern	0	0	0	0	0	0	70	15		0	0	0	0	0	0	80	14
Times N. R.	0	0	0	0	0	0	29	85		0	0	0	0	0	0	20	86

Gauss

KNN

	Gauss	KNN (N=3)
Resustitución	94.12	97.75
Cross-Validation	94.00	97.25
L.O.O.	92.87	94.87

Tabla 4.19. Resumen de probabilidades medias para el caso de 6 puntos, 135° rotación.

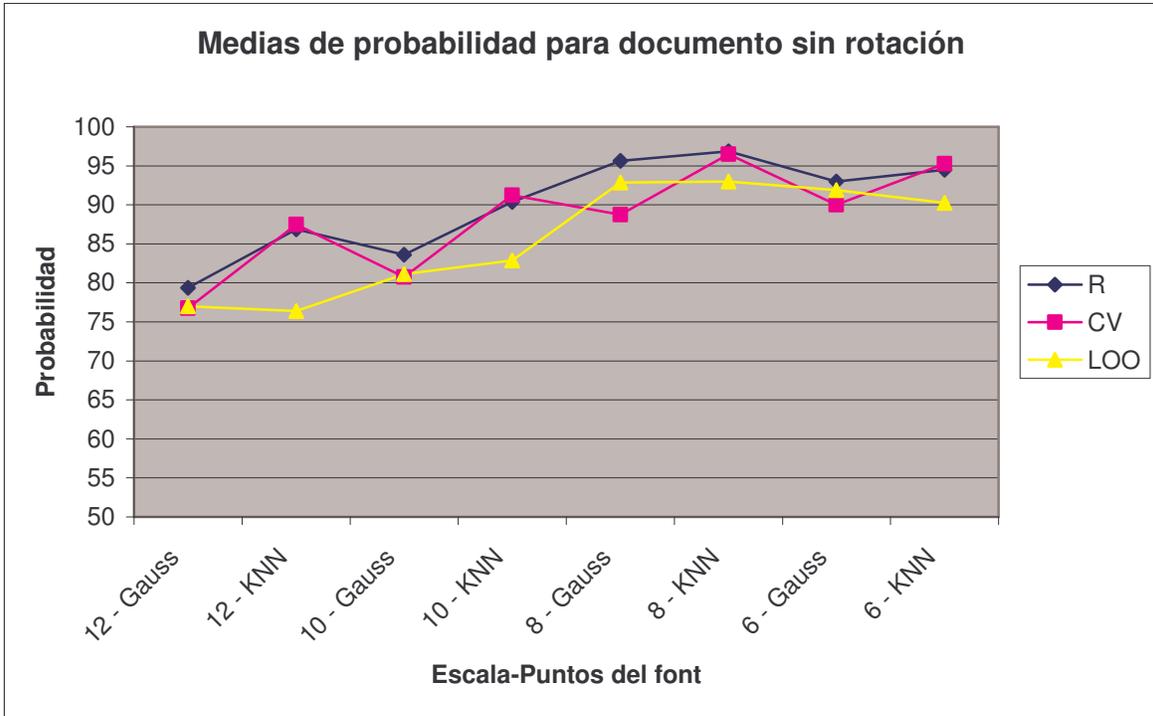
4.2.17. Resumiendo.

En la tabla 4.20 se muestra el resumen de todas las pruebas hechas en estos casos. Se puede apreciar que el mínimo de probabilidad se alcanza en la prueba de Cross Validation (CV), ahí recordemos que solo se entrena a los clasificadores con la mitad de la base de datos, mientras que el estimador que quizá más nos interese es el de Leave-One-Out (LOO), el cual permanece en su mayoría en valores intermedios entre el máximo y el mínimo.

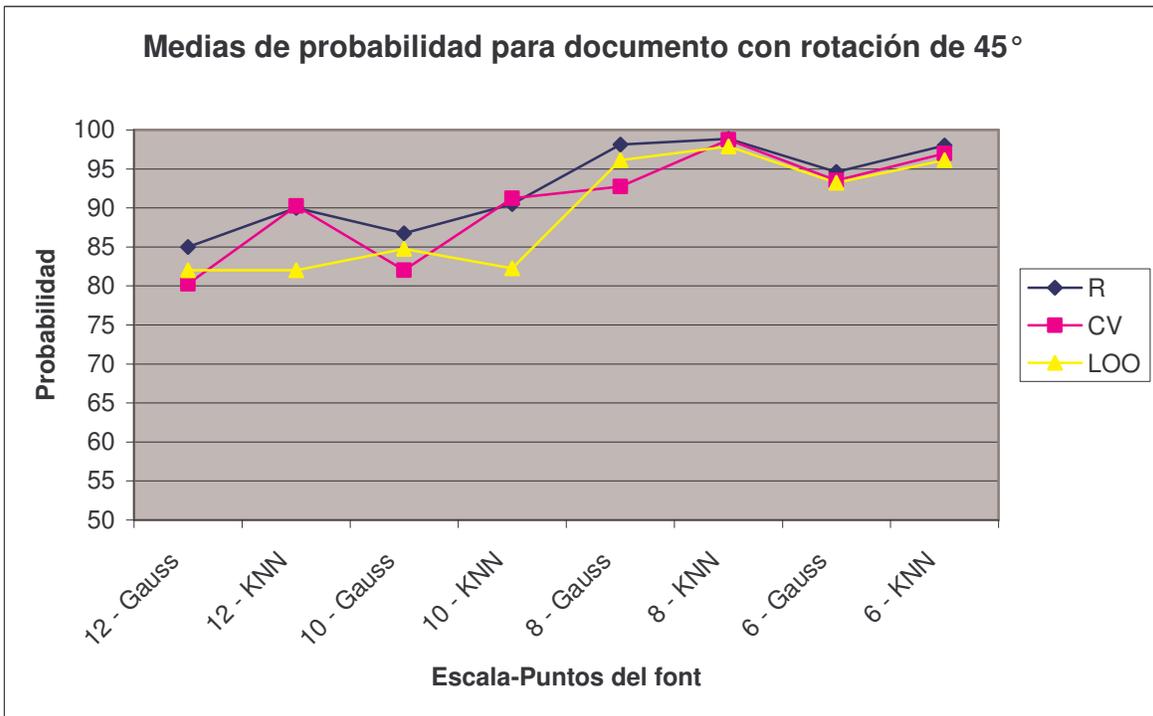
En las Gráficas 4.1 a 4.4 se puede apreciar los niveles de probabilidad, en general tenemos un mínimo global de 76.28, para el caso LOO a 12 pts., y el máximo global de 99.25 % para el caso de Resustitución, el método de validación más idealizado, para el caso de rotación de 90° a 8 pts de escala.

Rotación	0°			45°			90°			135°		
	R	CV	LOO	R	CV	LOO	R	CV	LOO	R	CV	LOO
12 - Gauss	79.38	76.75	77	85	80.25	82	82	76.75	79.5	84.13	79.5	82
12 - KNN	86.88	87.5	76.38	90	90.25	82	89.13	90	80.38	89.38	89.5	81
10 - Gauss	83.63	80.75	81.13	86.75	82	84.75	89.63	87.5	87.13	85	81.5	82.13
10 - KNN	90.38	91.25	82.88	90.5	91.25	82.25	90.88	93.5	83.5	91	91.75	82.5
8 - Gauss	95.63	88.75	92.88	98.13	92.75	96.13	95.38	93	93.25	96.88	95.25	95.13
8 - KNN	96.88	96.5	93	98.88	98.75	97.88	97.5	96.5	95.13	99.25	99	97.25
6 - Gauss	93	90	91.88	94.63	93.5	93.25	89.5	88	87.75	94.13	94	92.88
6 - KNN	94.5	95.25	90.25	98	97	96.13	93.88	95.75	89.38	97.75	97.25	94.88
Máximo	96.88			98.88			97.5			99.25		
Mínimo	76.38			80.25			76.75			79.5		

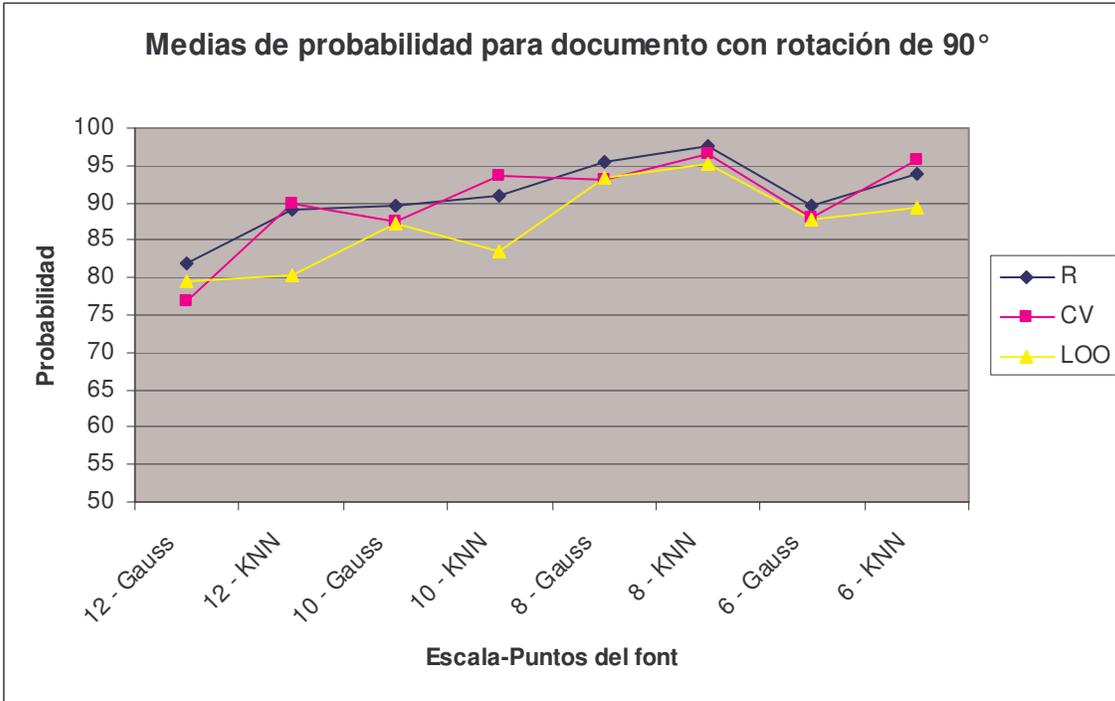
Tabla 4.20. Resumen global de probabilidades medias



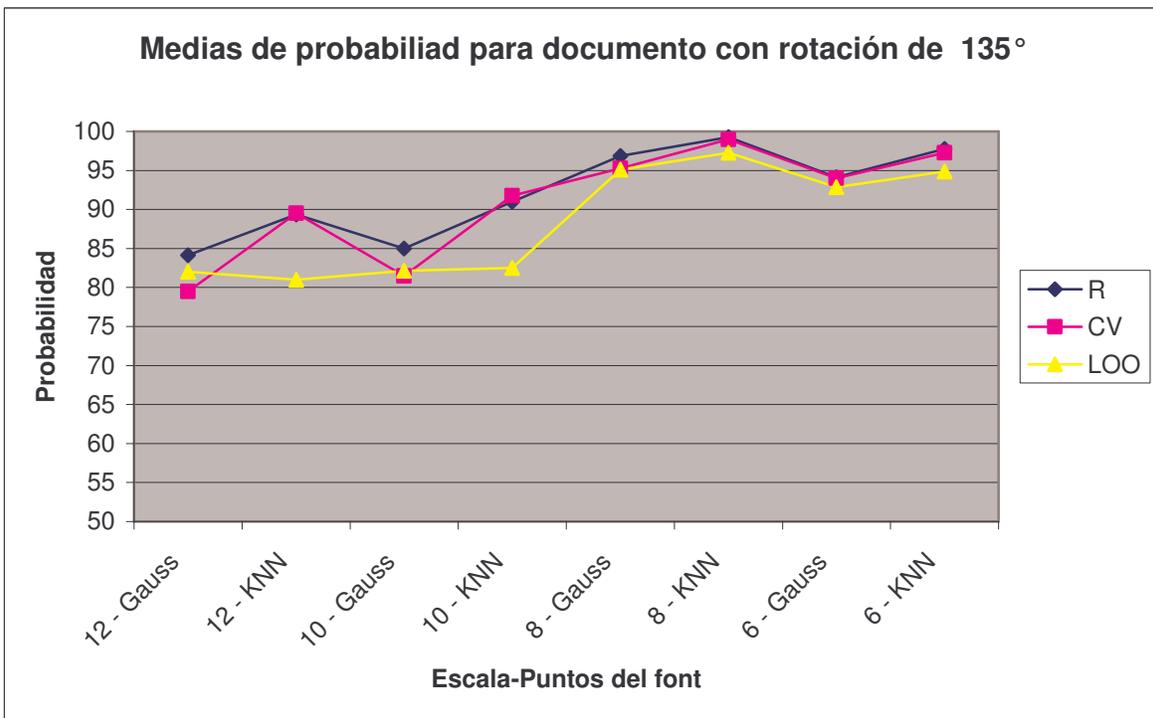
Gráfica 4.1.



Gráfica 4.2.



Gráfica 4.3.



Gráfica 4.4.

4.3. Validación por cálculo de matriz de confusión para el caso de letra itálica, negrita y combinación de itálica con negrita.

Los siguientes experimentos se hacen sobre los siguientes casos de tipos de font, tamaño en puntos y rotaciones, como se muestra en la Tabla 4.21. En estos casos el tipo de letra es con el efecto itálica (*i*), negrita (**b**), y la combinación de ambos efectos, itálica con negrita (*bi*). En la Figura 4.17 se ponen unas muestras de las clases con los efectos sobre el font.

N°	Font	Escala				Rotación				Tipo:		
		12	10	8	6	0°	45°	90°	135°	i	b	bi
1	Courier New	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻
2	Arial	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻
3	Bookman Old Style	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻
4	Franklin Ghotic Médium	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻
5	Comic Sans MS	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻
6	Impact	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻
7	Modern	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻
8	Times New Roman	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻	↻

Tabla 4.21. Todos los casos considerados en las pruebas, por cada escala se toman las 4 medidas de rotación (e.g. 12 pts. En 0°, 45°, 90° y 135°), y 3 efectos (itálica, negrita y combinación de itálica con negrita).

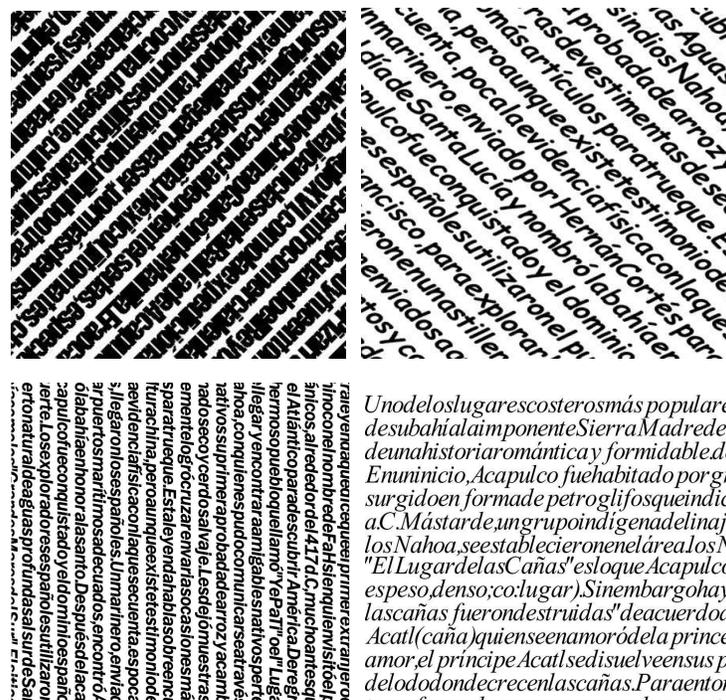


Figura 4.17. Muestra de unas clases de texto preprocesada sin distorsión, a 0°, 45°, 90° y 135° respectivamente, con efectos itálica, negrita y combinación de ambas.

A partir de la Tabla 4.21, podemos ver que tenemos un total de:

$$4 \text{ escalas} * 4 \text{ rotaciones} * 3 \text{ tipos} * 8 \text{ fonts} = 384 \text{ casos}$$

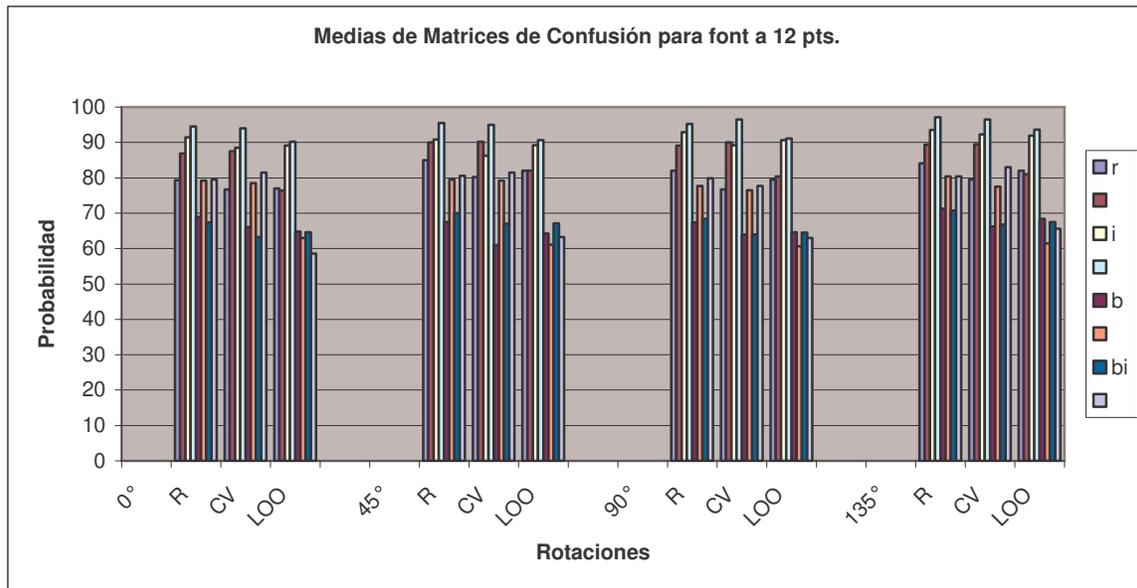
y de estos a mostrar 6 tablas (un par por tipo de validación de matriz de confusión). Para resumir presentamos en concentrado las probabilidades medias por cada caso, en cada tabla se marca el valor máximo y el mínimo. Para mayor detalle, se pueden consultar cada tabla por cada caso en la media de disco compacto que se anexa.

A continuación de cada tabla de concentrados se muestran los datos en gráfica para una mejor interpretación de los mismos.

4.3.1. Caso: Font a 12 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.

12 pts.		Efecto							
0°		r		i		b		bi	
<i>Medias de Mat. de Conf.</i>		Gauss	KNN	Gauss	KNN	Gauss	KNN	Gauss	KNN
R		79.37	86.87	91.50	94.50	68.87	79.25	67.37	79.50
CV		76.75	87.50	88.50	94.00	66.00	78.50	63.25	81.50
LOO		77.00	76.37	89.12	90.25	64.75	63.00	64.62	58.62
45°		r		i		b		bi	
R		85.00	90.00	90.88	95.50	67.50	79.50	70.00	80.63
CV		80.25	90.25	86.25	95.00	61.00	79.25	67.00	81.50
LOO		82.00	82.00	89.25	90.63	64.25	61.13	67.13	63.25
90°		r		i		b		bi	
R		82.00	89.12	92.88	95.25	67.38	77.75	68.38	79.88
CV		76.75	90.00	89.25	96.50	64.00	76.50	64.00	77.75
LOO		79.50	80.37	90.63	91.13	64.63	60.63	64.50	63.00
135°		r		i		b		bi	
R		84.12	89.37	93.50	97.13	71.25	80.38	70.75	80.38
CV		79.50	89.50	92.25	96.50	66.25	77.50	66.75	83.00
LOO		82.00	81.00	91.88	93.63	68.38	61.50	67.50	65.63

Tabla 4.22. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 12 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.

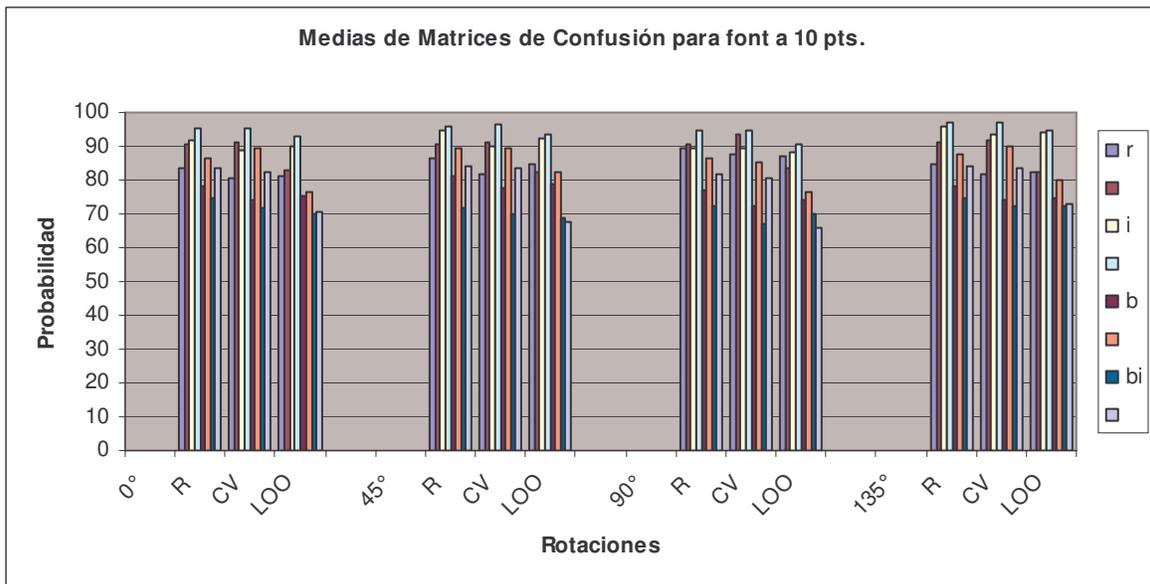


Gráfica 4.5.

4.3.2. Caso: Font a 10 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.

10 pts.								
Medias de Mat. de Conf.	Gauss	KNN	Gauss	KNN	Gauss	KNN	Gauss	KNN
0°	r		i		b		bi	
R	83.62	90.37	91.88	95.50	78.13	86.63	74.75	83.25
CV	80.75	91.25	88.75	95.50	74.00	89.25	71.75	82.50
LOO	81.12	82.87	89.88	93.13	75.13	76.63	70.25	70.37
45°	r		i		b		bi	
R	86.75	90.50	94.63	95.88	81.13	89.63	71.88	84.38
CV	82.00	91.25	90.00	96.50	77.75	89.25	70.25	83.25
LOO	84.75	82.25	92.50	93.63	78.88	82.38	69.00	67.63
90°	r		i		b		bi	
R	89.62	90.87	89.38	94.75	77.25	86.75	72.25	81.75
CV	87.50	93.50	89.25	95.00	72.50	85.25	67.25	80.75
LOO	87.12	83.50	88.00	90.38	74.38	76.38	69.88	65.88
135°	r		i		b		bi	
R	85.00	91.00	96.00	97.25	78.25	87.63	74.63	84.38
CV	81.50	91.75	93.50	97.25	74.25	89.75	72.25	83.75
LOO	82.125	82.50	94.25	94.75	74.88	79.75	72.50	72.88

Tabla 4.23. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 10 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.

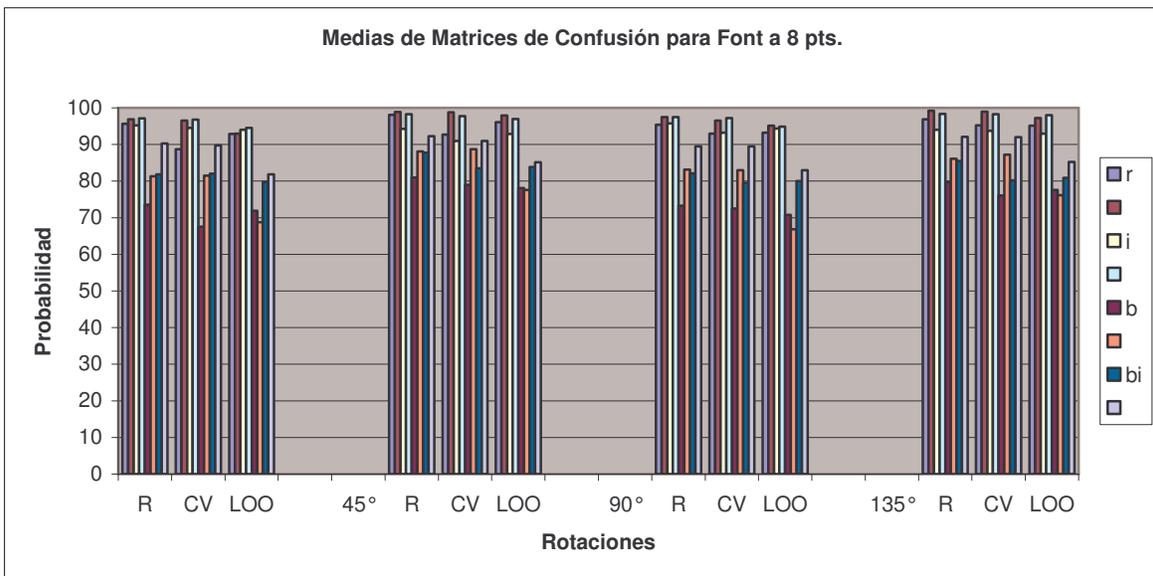


Gráfica 4.6.

4.3.3. Caso: Font a 8 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.

8 pts.								
Medias de Mat. de Conf.	Gauss	KNN	Gauss	KNN	Gauss	KNN	Gauss	KNN
0°	r		i		b		bi	
R	95.62	96.87	95.25	97.13	73.50	81.38	81.88	90.25
CV	88.75	96.50	94.50	96.75	67.50	81.50	82.00	89.75
LOO	92.875	93.00	94.00	94.50	71.88	68.75	79.75	81.88
45°	r		i		b		bi	
R	98.12	98.87	94.25	98.25	81.00	88.13	87.75	92.25
CV	92.75	98.75	91.00	97.75	79.00	88.75	83.50	91.00
LOO	96.12	97.875	92.88	97.00	78.13	77.63	83.88	85.13
90°	r		i		b		bi	
R	95.37	97.50	95.75	97.50	73.25	83.13	82.13	89.50
CV	93.00	96.50	93.25	97.25	72.50	83.00	79.50	89.50
LOO	93.25	95.12	94.38	94.88	70.75	66.88	80.00	83.00
135°	r		i		b		bi	
R	96.87	99.25	94.00	98.38	79.75	86.13	85.50	92.13
CV	95.25	99.00	93.75	98.25	76.00	87.25	80.25	92.00
LOO	95.12	97.25	93.00	98.00	77.63	76.13	80.88	85.25

Tabla 4.24. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 8 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.

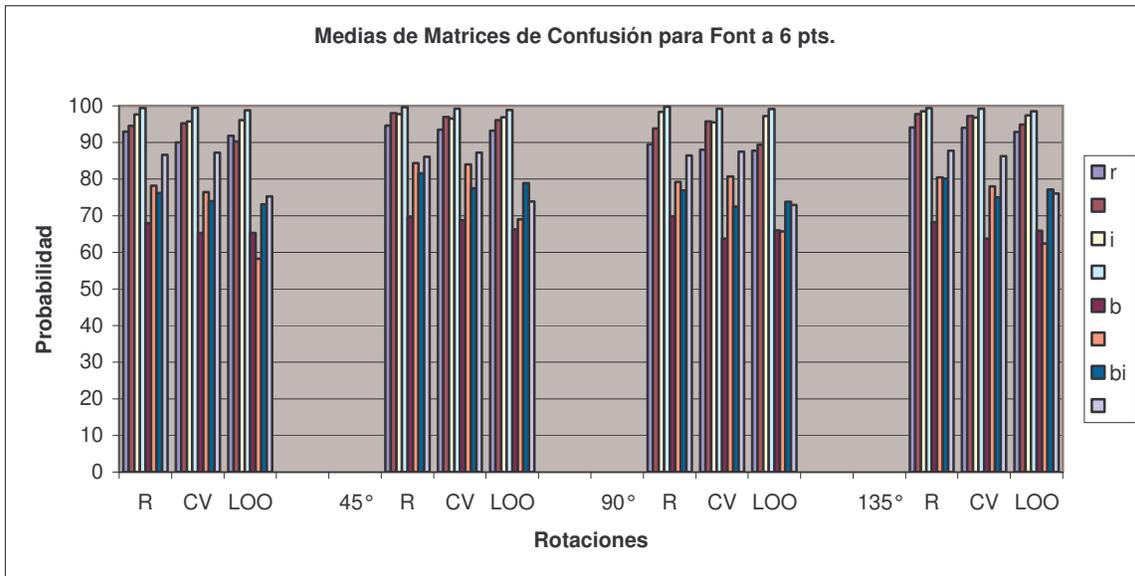


Gráfica 4.7.

4.3.4. Caso: Font a 6 pts, rotaciones 0°, 45°, 90° y 135°; efectos: Negrita, Itálica y combinación.

6 pts.								
Medias de Mat. de Conf.	Gauss	KNN	Gauss	KNN	Gauss	KNN	Gauss	KNN
0°	r		i		b		bi	
R	93.00	94.50	97.63	99.38	68.00	78.25	76.25	86.63
CV	90.00	95.25	95.75	99.50	65.25	76.50	74.00	87.25
LOO	91.87	90.25	96.13	98.75	65.25	58.25	73.13	75.25
45°	r		i		b		bi	
R	94.62	98.00	97.75	99.63	69.50	84.38	81.63	86.13
CV	93.50	97.00	96.50	99.25	68.75	84.00	77.50	87.25
LOO	93.25	96.12	96.88	98.88	66.25	69.00	78.88	73.88
90°	R		i		b		bi	
R	89.50	93.87	98.38	99.75	69.75	79.25	76.88	86.50
CV	88.00	95.75	95.50	99.25	63.75	80.75	72.50	87.50
LOO	87.75	89.37	97.25	99.13	66.00	65.75	73.75	72.88
135°	R		i		b		bi	
R	94.12	97.75	98.50	99.38	68.25	80.50	80.13	87.75
CV	94.00	97.25	96.75	99.25	63.75	78.00	75.00	86.25
LOO	92.87	94.87	97.38	98.50	65.88	62.38	77.13	76.00

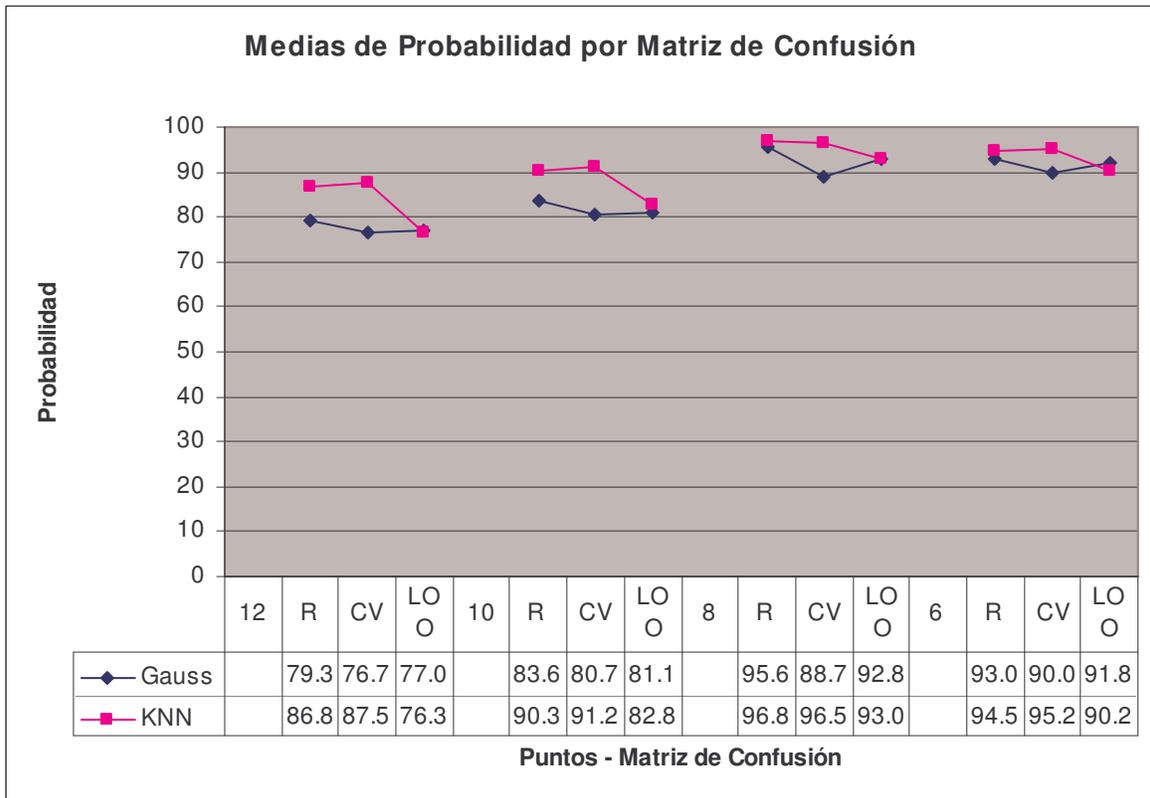
Tabla 4.25. Concentrado de medias de matrices de confusión para ambos clasificadores para el caso 6 pts / rotaciones: 0°, 45°, 90°, 135° / efectos: i, b, bi.



Gráfica 4.8.

4.3.5. Resumiendo los casos de Itálica, Negrita y combinación de ambas.

Como podemos apreciar de las Gráficas 4.5 a 4.8, ahora resumiendo en la Gráfica 4.9, para font regular, conservando este tamaño de ventana de prueba, de 256 píxeles, se obtiene un mayor nivel de acertividad en las diferentes validaciones a un tamaño de font menor, con 8 y 6 puntos, por arriba del 90%, para el clasificador basado en KNN, y en general para todos los casos, esta por arriba del proceso Gaussiano.



Gráfica 4.9.

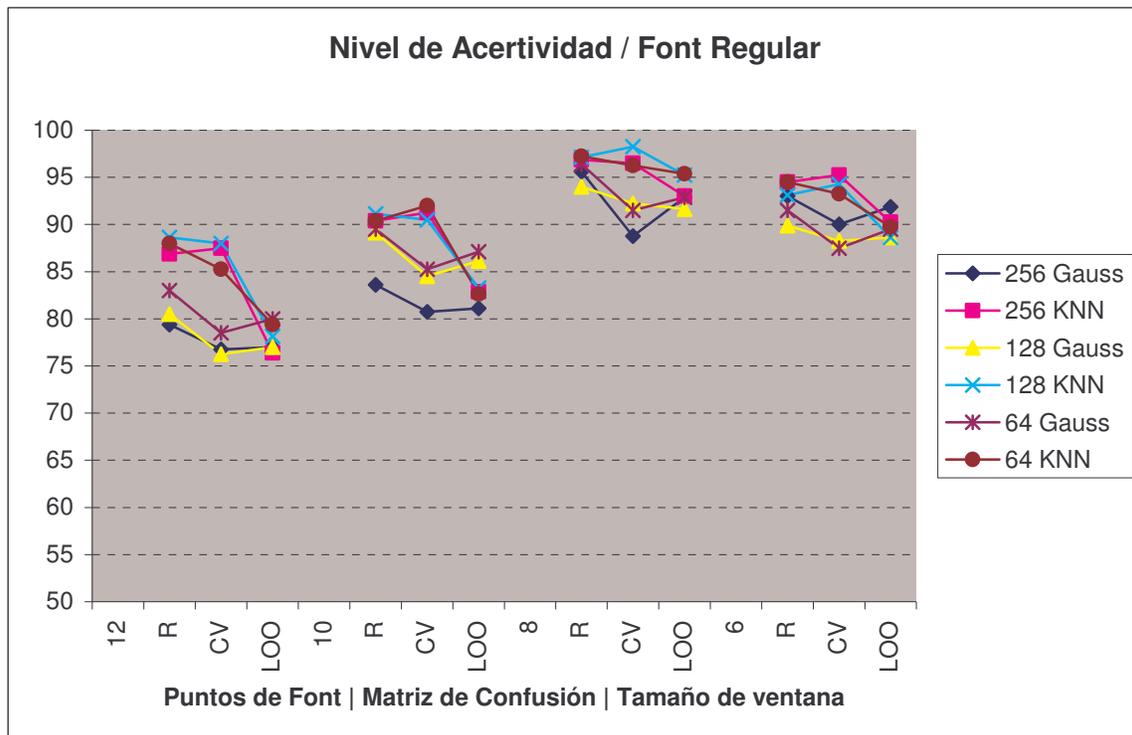
4.4. Experimentos con diferente tamaño de ventana de prueba.

En esta sección presentamos los resultados de experimentar con diferentes tamaños de ventana de prueba, en el barrido aleatorio sobre las clases, para conformar nuestra base de datos de aprendizaje.

Primero mostraremos los resultados de niveles de acertividad obtenidos con los tres diferentes tamaños que experimentamos (256, 128 y 64 píxeles), los efectos sobre el font (Regular, Itálica, Negrita, Itálica + negrita), y los métodos de validación usados (Matrices de confusión: R, CV, LOO).

Tamaño ventana:	256		128		64	
	Gauss	KNN	Gauss	KNN	Gauss	KNN
Puntos: 12						
R	79.37	86.87	80.50	88.63	83.00	88.00
CV	76.75	87.50	76.25	88.00	78.50	85.25
LOO	77.00	76.37	77.00	78.13	80.00	79.38
Puntos: 10						
R	83.62	90.37	89.13	91.13	89.50	90.38
CV	80.75	91.25	84.50	90.50	85.25	92.00
LOO	81.12	82.87	86.13	83.25	87.13	82.63
Puntos: 8						
R	95.62	96.87	94.00	97.13	96.38	97.25
CV	88.75	96.50	92.25	98.25	91.50	96.25
LOO	92.88	93.00	91.63	95.25	92.88	95.38
Puntos: 6						
R	93.00	94.50	89.88	93.13	91.50	94.50
CV	90.00	95.25	88.25	94.25	87.50	93.25
LOO	91.87	90.25	88.63	88.63	89.50	89.75

Tabla 4.26. Medias de probabilidad, con rotación 0°, en font Regular

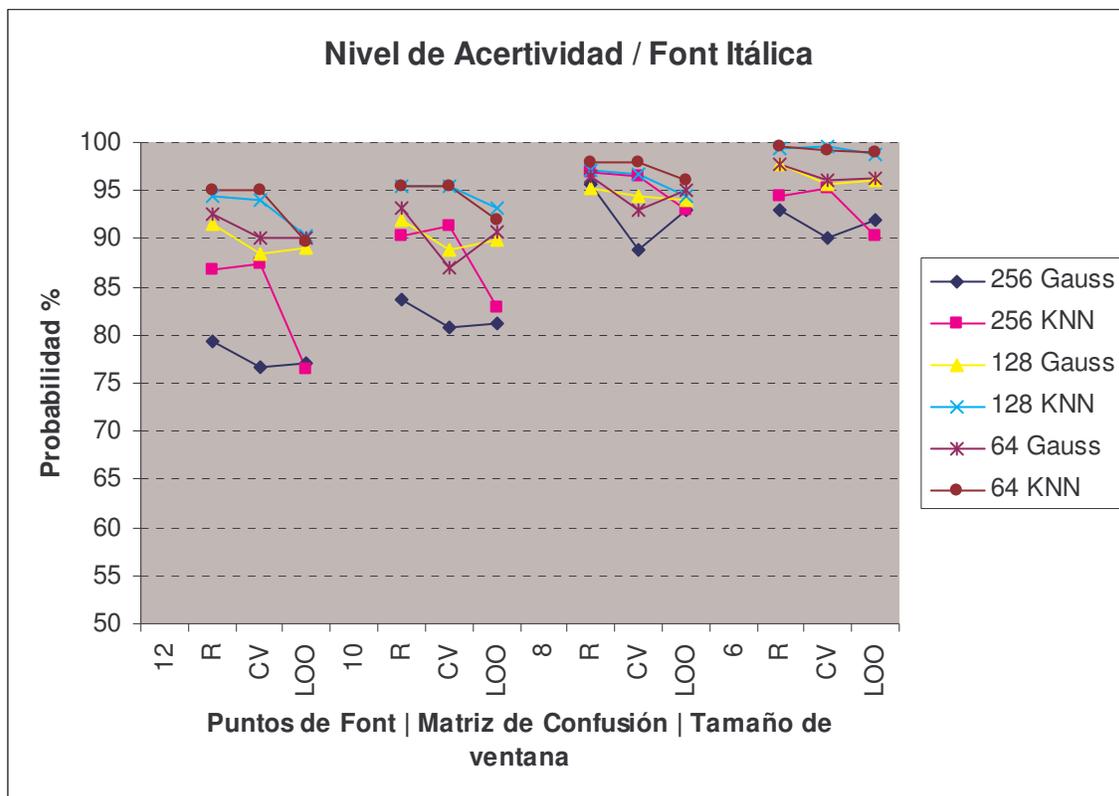


Gráfica 4.10.

Tamaño ventana:	256	128	64
-----------------	-----	-----	----

	Gauss	KNN	Gauss	KNN	Gauss	KNN
Puntos: 12						
R	79.37	86.87	91.50	94.50	92.63	95.00
CV	76.75	87.50	88.50	94.00	90.00	95.00
LOO	77.00	76.37	89.13	90.25	90.00	89.63
Puntos: 10						
R	83.62	90.37	91.88	95.50	93.25	95.50
CV	80.75	91.25	88.75	95.50	87.00	95.50
LOO	81.12	82.87	89.88	93.13	90.63	91.88
Puntos: 8						
R	95.62	96.87	95.25	97.13	96.50	97.88
CV	88.75	96.50	94.50	96.75	93.00	98.00
LOO	92.88	93.00	94.00	94.50	95.00	96.13
Puntos: 6						
R	93.00	94.50	97.63	99.38	97.75	99.50
CV	90.00	95.25	95.75	99.50	96.00	99.25
LOO	91.87	90.25	96.13	98.75	96.38	98.88

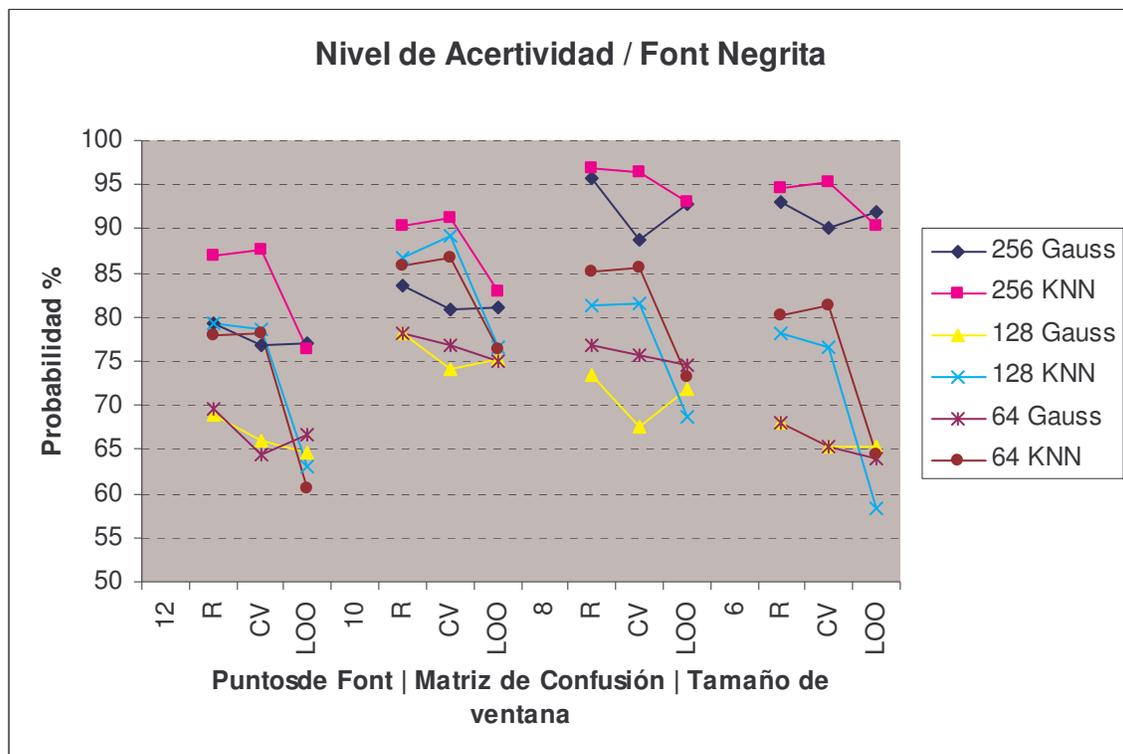
Tabla 4.27. Medias de probabilidad, con rotación 0°, en font Itálica



Gráfica 4.11.

Tamaño ventana:	256		128		64	
	Gauss	KNN	Gauss	KNN	Gauss	KNN
Puntos: 12						
R	79.37	86.87	68.88	79.25	69.50	77.88
CV	76.75	87.50	66.00	78.50	64.50	78.25
LOO	77.00	76.37	64.75	63.00	66.63	60.63
Puntos: 10						
R	83.62	90.37	78.13	86.63	78.13	85.75
CV	80.75	91.25	74.00	89.25	76.75	86.75
LOO	81.12	82.87	75.13	76.63	75.00	76.38
Puntos: 8						
R	95.62	96.87	73.50	81.38	76.75	85.13
CV	88.75	96.50	67.50	81.50	75.75	85.50
LOO	92.88	93.00	71.88	68.75	74.63	73.13
Puntos: 6						
R	93.00	94.50	68.00	78.25	68.00	80.13
CV	90.00	95.25	65.25	76.50	65.25	81.25
LOO	91.87	90.25	65.25	58.25	63.88	64.38

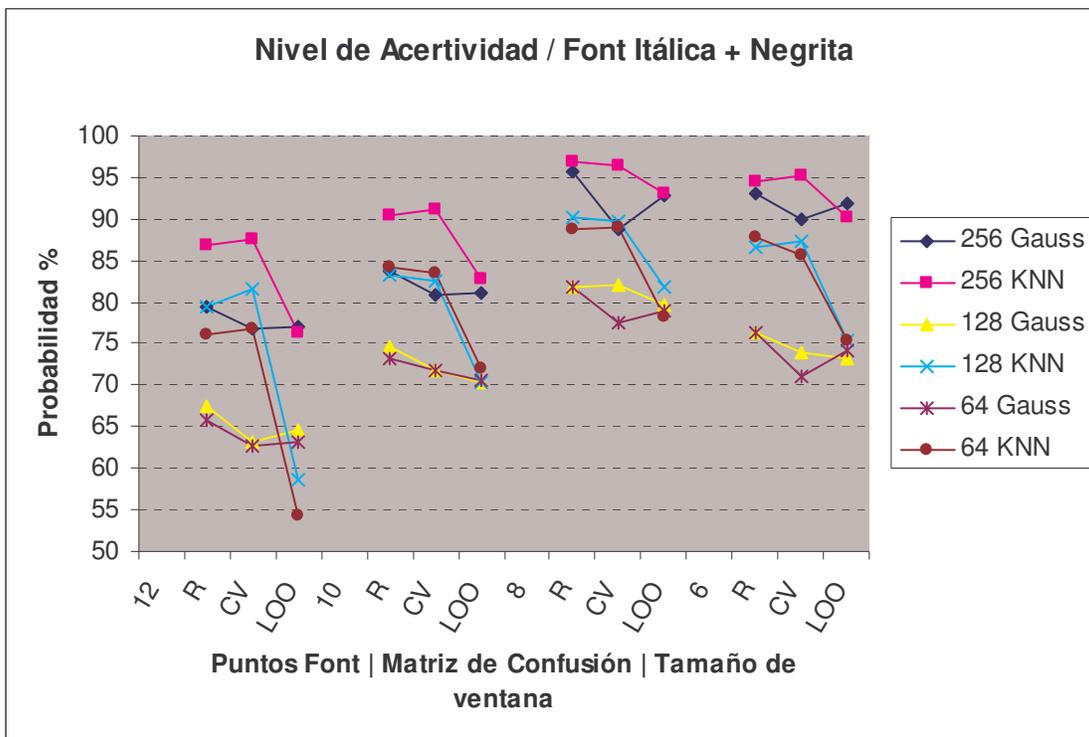
Tabla 4.28. Medias de probabilidad, con rotación 0°, en font Negrita



Gráfica 4.12.

Tamaño ventana:	256		128		64	
	Gauss	KNN	Gauss	KNN	Gauss	KNN
Puntos: 12						
R	79.37	86.87	67.38	79.50	65.88	76.13
CV	76.75	87.50	63.25	81.50	62.75	76.75
LOO	77.00	76.37	64.63	58.63	63.25	54.25
Puntos: 10						
R	83.62	90.37	74.75	83.25	73.13	84.25
CV	80.75	91.25	71.75	82.50	71.75	83.50
LOO	81.12	82.87	70.25	70.38	70.50	72.00
Puntos: 8						
R	95.62	96.87	81.88	90.25	81.75	88.75
CV	88.75	96.50	82.00	89.75	77.50	89.00
LOO	92.88	93.00	79.75	81.88	78.88	78.25
Puntos: 6						
R	93.00	94.50	76.25	86.63	76.38	87.75
CV	90.00	95.25	74.00	87.25	71.00	85.75
LOO	91.87	90.25	73.13	75.25	74.25	75.38

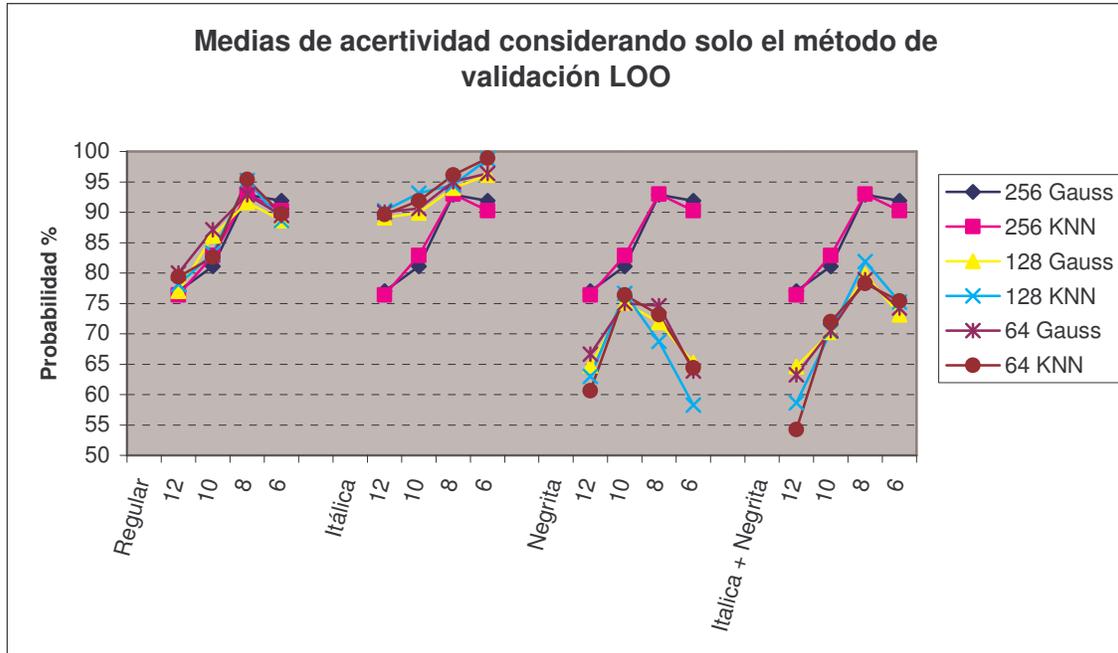
Tabla 4.29. Medias de probabilidad, con rotación 0°, en font Itálica con Negrita



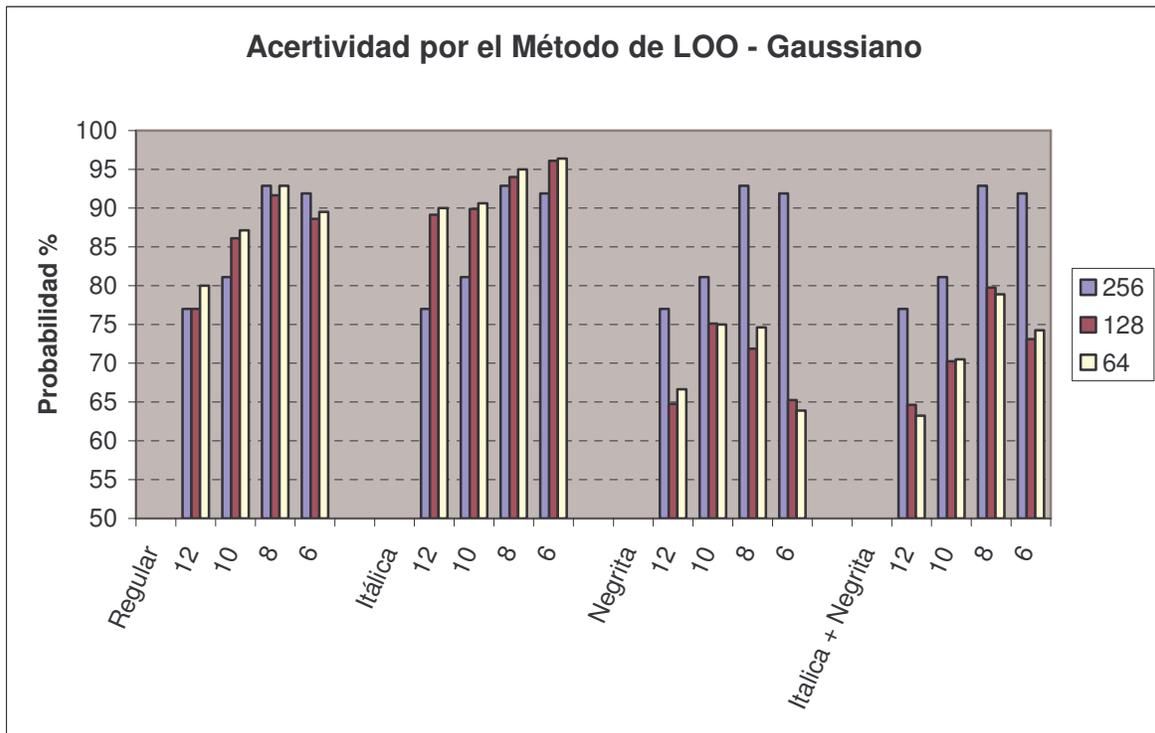
En la Tabla 4.30, con sus gráficas respectivas, mostramos únicamente los resultados para el método de validación de Leave-One-Out (LOO) por considerarse el más realista para la aplicación.

Tamaño ventana:	256		128		64	
	Gauss	KNN	Gauss	KNN	Gauss	KNN
Regular						
12	77.00	76.37	77.00	78.13	80.00	79.38
10	81.12	82.87	86.13	83.25	87.13	82.63
8	92.88	93.00	91.63	95.25	92.88	95.38
6	91.87	90.25	88.63	88.63	89.50	89.75
Itálica						
12	77.00	76.37	89.13	90.25	90.00	89.63
10	81.12	82.87	89.88	93.13	90.63	91.88
8	92.88	93.00	94.00	94.50	95.00	96.13
6	91.87	90.25	96.13	98.75	96.38	98.88
Negrita						
12	77.00	76.37	64.75	63.00	66.63	60.63
10	81.12	82.87	75.13	76.63	75.00	76.38
8	92.88	93.00	71.88	68.75	74.63	73.13
6	91.87	90.25	65.25	58.25	63.88	64.38
Itálica + Negrita						
12	77.00	76.37	64.63	58.63	63.25	54.25
10	81.12	82.87	70.25	70.38	70.50	72.00
8	92.88	93.00	79.75	81.88	78.88	78.25
6	91.87	90.25	73.13	75.25	74.25	75.38

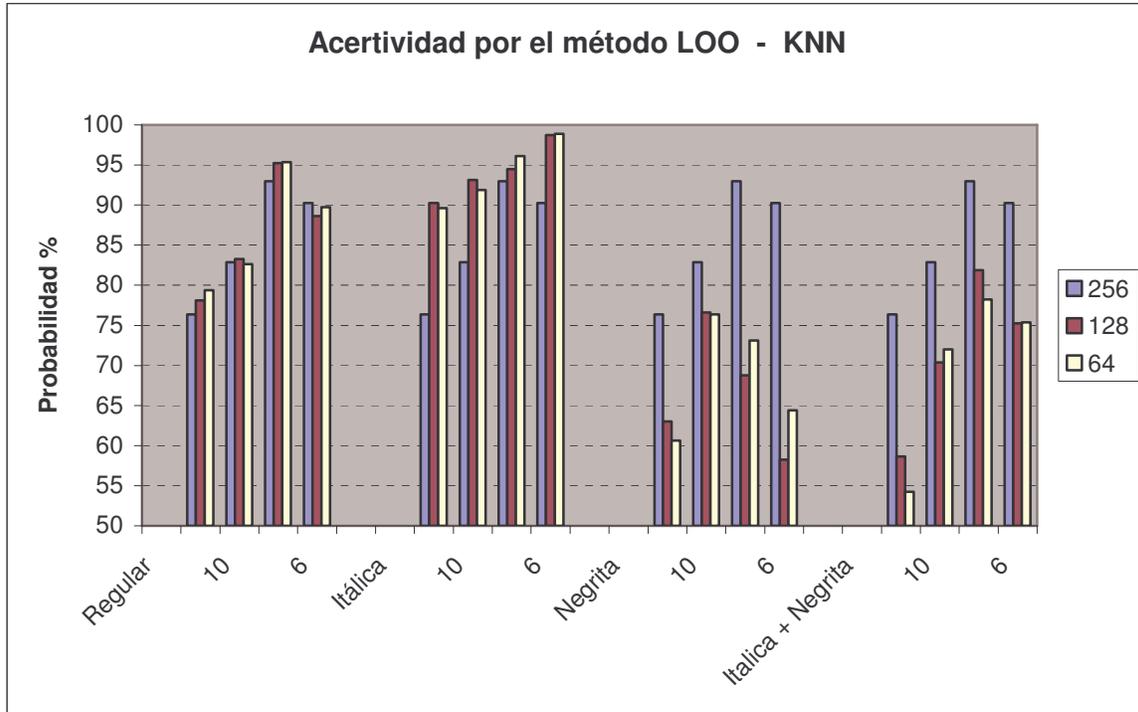
Tabla 4.30. Medias de probabilidad para el método LOO, con rotación 0°



Gráfica 4.14. Los datos de la Tabla 4.30.



Gráfica 4.15. De los datos de la Tabla 4.30, solo las medias de probabilidad para el clasificador Gaussiano, por el método de validación de LOO.



Gráfica 4.16. De los datos de la Tabla 4.30, solo las medias de probabilidad para el clasificador KNN, por el método de validación de LOO

Para tener una mejor comprensión visual del tamaño de la ventana, en la Figura 4.18 se muestran unas ventanas con tamaños aproximados relativos, conservando proporción entre ellas, las muestras de las ventanas son sobre clases de fonts a 10 puntos de tamaño todas, a 300 DPI de resolución.



Figura 4.18. Muestras de ventanas de prueba a 256, 128 y 64 píxeles, de arriba a bajo sucesivamente, con diferentes efectos sobre el font de 10 puntos, a 300 DPI.

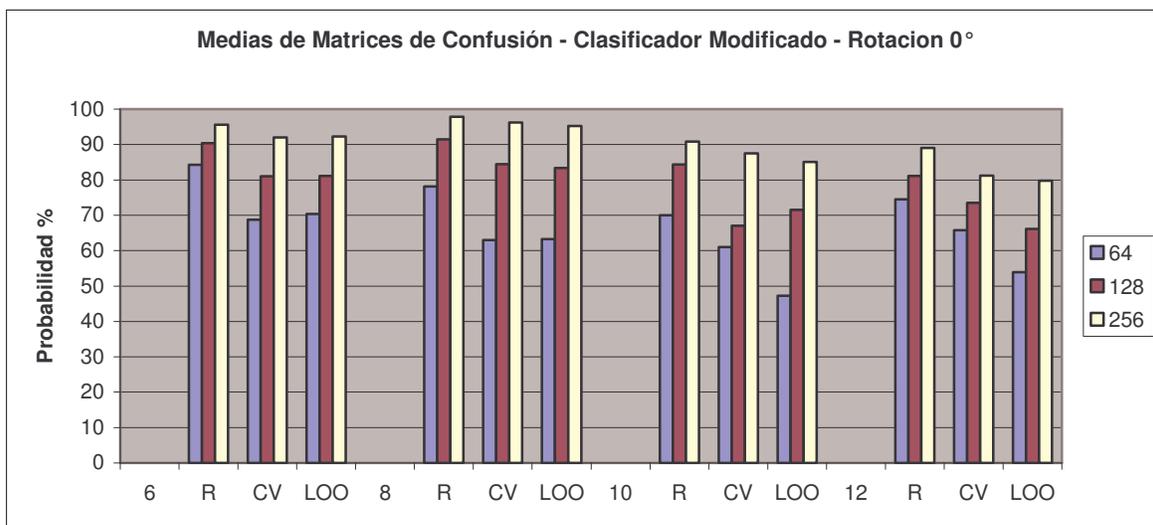
4.5. Pruebas con el clasificador modificado.

4.4.1. 4.5.1. Pruebas sobre el font Regular.

Ahora a partir de los resultados anteriores, podemos construir nuestro clasificador modificado, como una combinación de ambos métodos, Gaussiano y KNN, y al final de la clasificación por ambos métodos, tomamos el que arroje la mayor probabilidad. Los resultados son como eran de esperarse. Se muestran en las Tablas y Gráficas siguientes en resumen.

	64	128	256
6			
R	84.25	90.38	95.63
CV	68.75	81.00	92.00
LOO	70.38	81.13	92.25
8			
R	78.13	91.50	97.88
CV	63.00	84.50	96.25
LOO	63.25	83.38	95.25
10			
R	70.00	84.38	90.88
CV	61.00	67.00	87.50
LOO	47.25	71.50	85.13
12			
R	74.50	81.13	89.00
CV	65.75	73.50	81.25
LOO	53.88	66.13	79.75

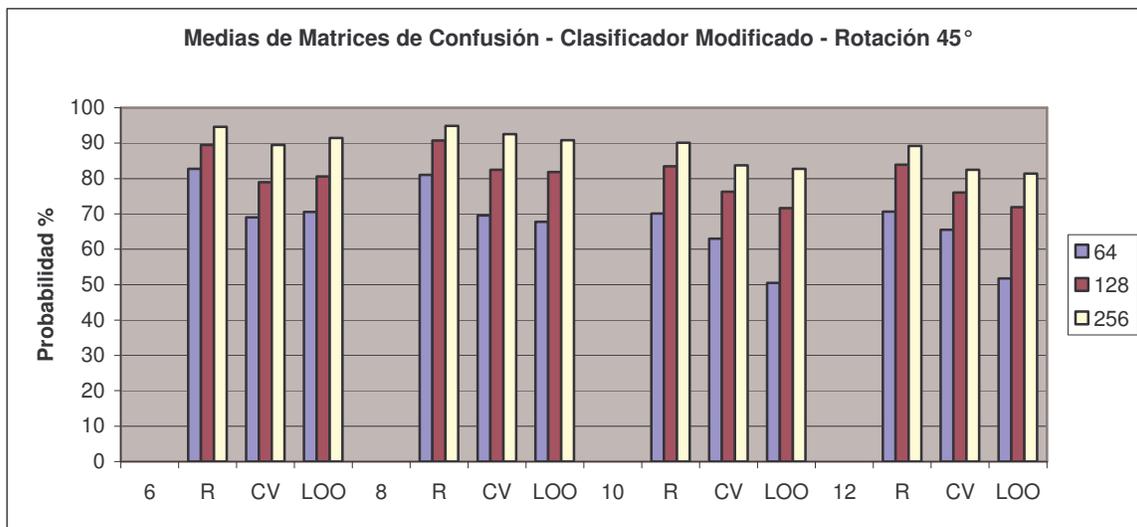
Tabla 4.30. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 0°.



Grafica 4.17. Medias probabilidades, clasificador modificado, Font regular, Rotación 0°.

	64	128	256
6			
R	82.75	89.50	94.63
CV	69.00	79.00	89.50
LOO	70.50	80.63	91.50
8			
R	81.00	90.75	94.88
CV	69.50	82.50	92.50
LOO	67.75	81.88	90.88
10			
R	70.13	83.50	90.13
CV	63.00	76.25	83.75
LOO	50.50	71.63	82.75
12			
R	70.63	83.88	89.25
CV	65.50	76.00	82.50
LOO	51.75	71.88	81.38

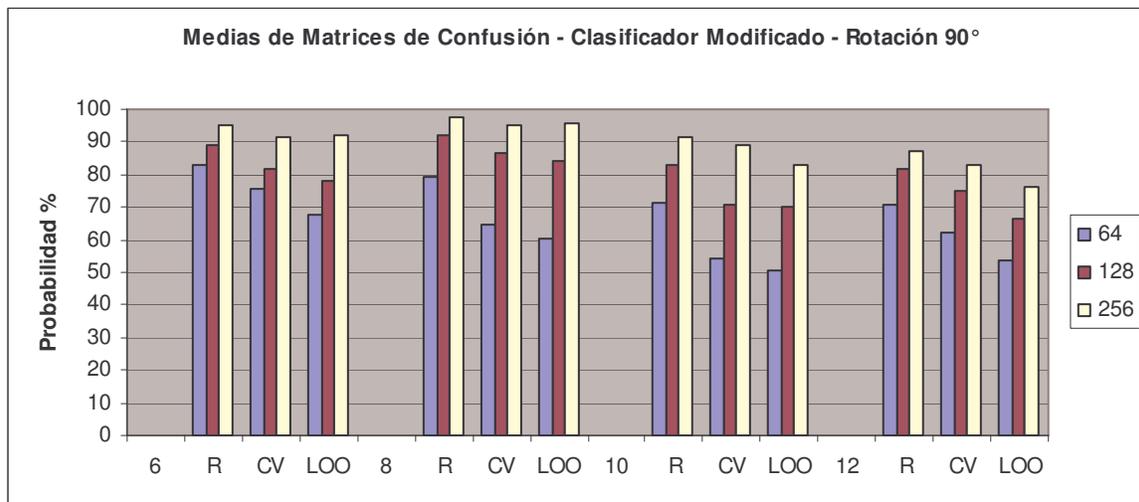
Tabla 4.31. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 45°.



Grafica 4.18. Medias probabilidades, clasificador modificado, Font regular, Rotación 45°.

	64	128	256
6			
R	82.88	88.88	95.13
CV	75.75	81.75	91.25
LOO	67.75	78.13	92.38
8			
R	79.38	92.38	97.63
CV	64.50	86.50	95.25
LOO	60.25	84.25	95.63
10			
R	71.13	83.00	91.25
CV	54.00	70.75	89.25
LOO	50.88	70.38	83.00
12			
R	70.50	81.63	87.25
CV	62.50	75.25	83.00
LOO	53.75	66.50	76.25

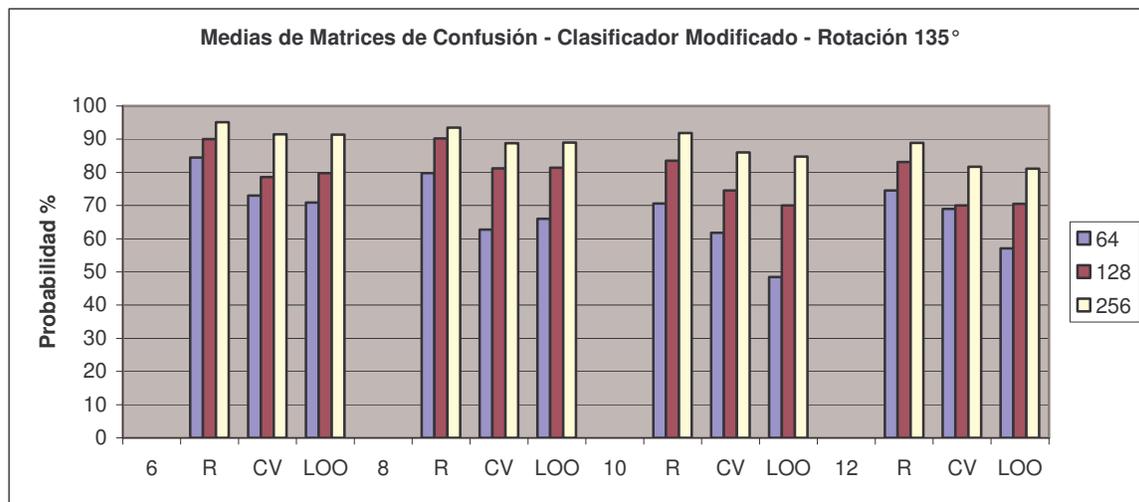
Tabla 4.32. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 90°.



Grafica 4.19. Medias probabilidades, clasificador modificado, Font regular, Rotación 90°.

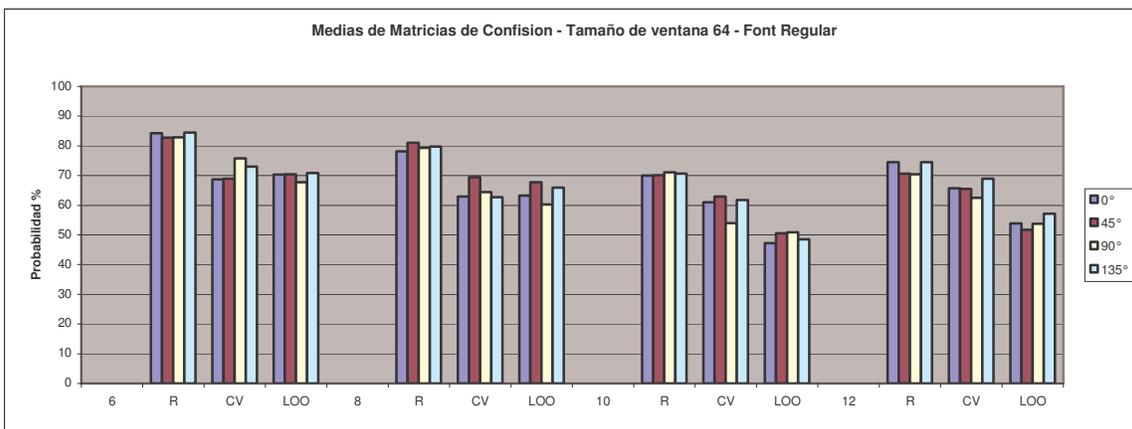
	64	128	256
6			
R	84.50	90.00	95.13
CV	73.00	78.50	91.50
LOO	70.88	79.75	91.38
8			
R	79.75	90.25	93.50
CV	62.75	81.25	88.75
LOO	66.00	81.38	89.00
10			
R	70.63	83.50	91.88
CV	61.75	74.50	86.00
LOO	48.50	70.00	84.75
12			
R	74.50	83.13	88.88
CV	69.00	70.00	81.75
LOO	57.13	70.50	81.13

Tabla 4.33. Tabla de medias probabilidades, clasificador modificado, Font regular, Rotación 135°.

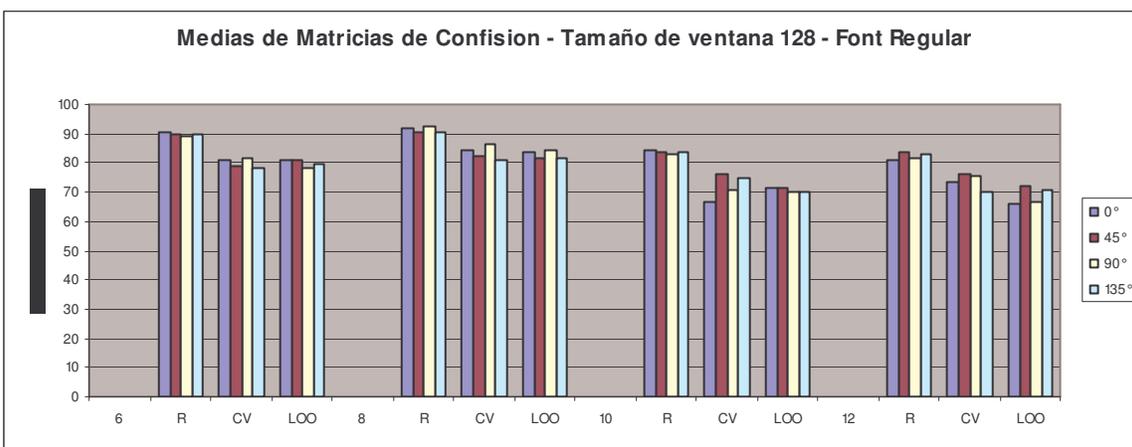


Grafica 4.20. Medias probabilidades, clasificador modificado, Font regular, Rotación 135°.

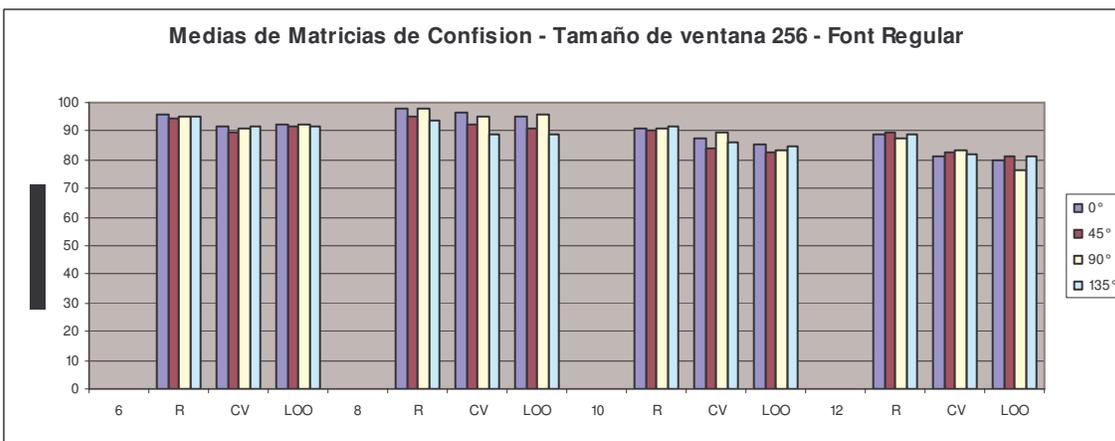
En esta serie de Tablas y Gráficas anteriores mostramos los datos dentro de una sola rotación, a continuación los condensamos por tamaño de ventana para tener otra apreciación.



Grafica 4.21. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones de 0°, 45°, 90° y 135°, font Regular.



Grafica 4.22. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones de 0°, 45°, 90° y 135°, font Regular.



Grafica 4.23. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones de 0°, 45°, 90° y 135°, font Regular.

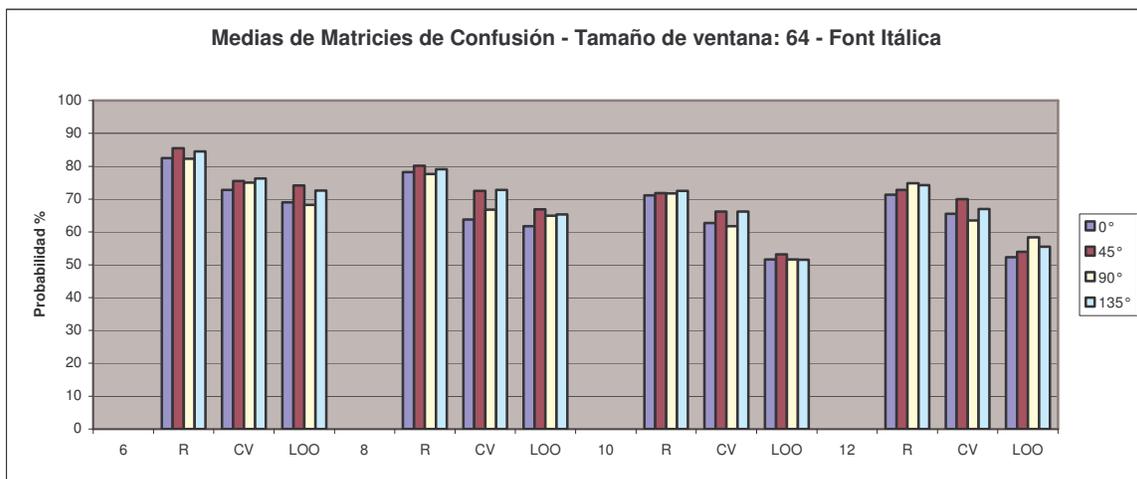
De la Grafica 4.23, con tamaño de ventana de 256 píxeles, validamos nuestra mejor opción, tomar este tamaño de ventana nos proporciona el mayor nivel de acertividad en nuestro clasificador que combina ambos procesos (Gaussiano y KNN), obteniendo niveles por arriba del 90% por el método de Leave-One-Out, para un tamaño de font de 6 y 8 puntos.

4.4.2. 4.5.2. Pruebas sobre los font Itálica, Negrita y Negrita con Itálica.

A continuación se muestran los resultados de los experimentos hechos sobre documentos con texto no regular (itálica, negrita, e itálica con negrita). Las Tablas y Gráficas se presentan concentradas por tamaño de ventana de prueba, con las diferentes rotaciones.

	0°	45°	90°	135°
6				
R	82.50	85.50	82.25	84.50
CV	72.75	75.50	75.00	76.25
LOO	69.00	74.13	68.25	72.63
8				
R	78.25	80.13	77.63	79.13
CV	63.75	72.50	66.75	72.75
LOO	61.75	66.88	65.00	65.38
10				
R	71.13	71.88	71.75	72.50
CV	62.75	66.25	61.75	66.25
LOO	51.63	53.13	51.63	51.50
12				
R	71.38	72.75	74.88	74.25
CV	65.50	70.00	63.50	67.00
LOO	52.25	53.88	58.38	55.50

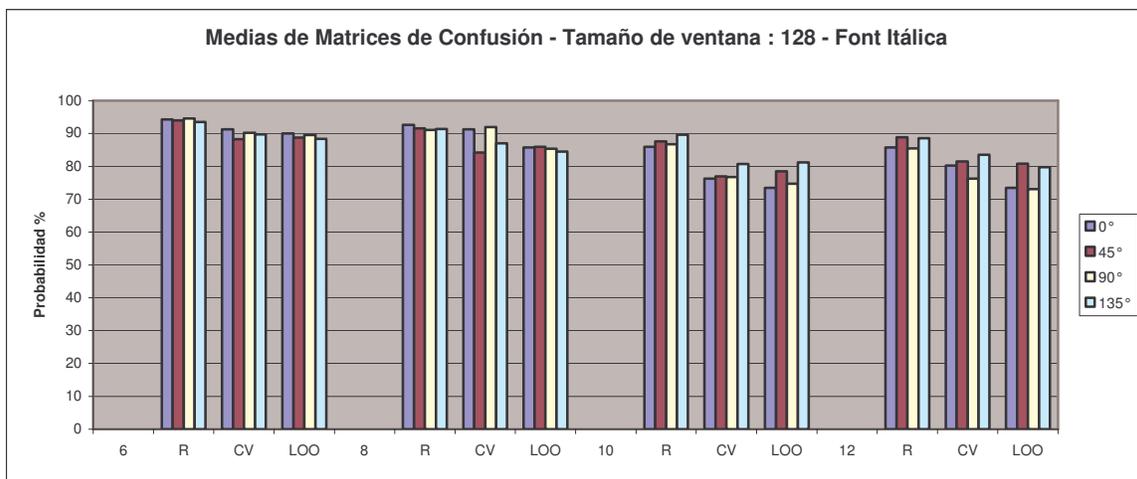
Tabla 4.34. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 64 píxeles, Font Itálica, Rotación 0°, 45°, 90°, 135°.



Grafica 4.24. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones 0°, 45°, 90° y 135°, Font Itálica

	0°	45°	90°	135°
6				
R	94.25	94.00	94.63	93.50
CV	91.25	88.25	90.25	89.75
LOO	90.00	88.75	89.50	88.38
8				
R	92.63	91.63	91.13	91.38
CV	91.25	84.25	92.00	87.00
LOO	85.75	86.00	85.38	84.50
10				
R	86.00	87.63	86.75	89.63
CV	76.25	77.00	76.75	80.75
LOO	73.50	78.50	74.75	81.25
12				
R	85.75	88.88	85.50	88.63
CV	80.25	81.50	76.25	83.50
LOO	73.50	80.88	73.13	79.75

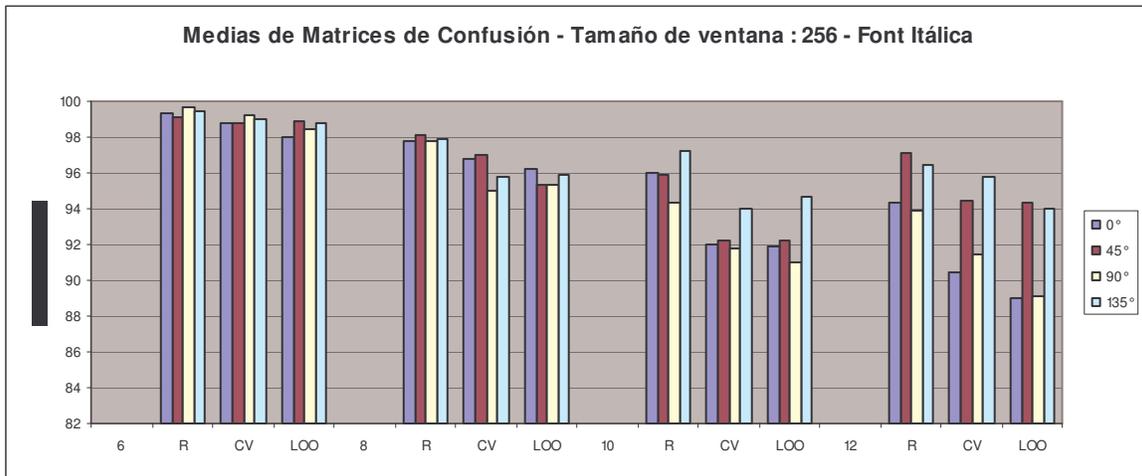
Tabla 4.35. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 128 píxeles, Font Itálica, Rotación 0°, 45°, 90°, 135°.



Grafica 4.25. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones 0°, 45°, 90° y 135°, Font Itálica

	0°	45°	90°	135°
6				
R	99.38	99.13	99.63	99.50
CV	98.75	98.75	99.25	99.00
LOO	98.00	98.88	98.50	98.75
8				
R	97.75	98.13	97.75	97.88
CV	96.75	97.00	95.00	95.75
LOO	96.25	95.38	95.38	95.88
10				
R	96.00	95.88	94.38	97.25
CV	92.00	92.25	91.75	94.00
LOO	91.88	92.25	91.00	94.63
12				
R	94.38	97.13	93.88	96.50
CV	90.50	94.50	91.50	95.75
LOO	89.00	94.38	89.13	94.00

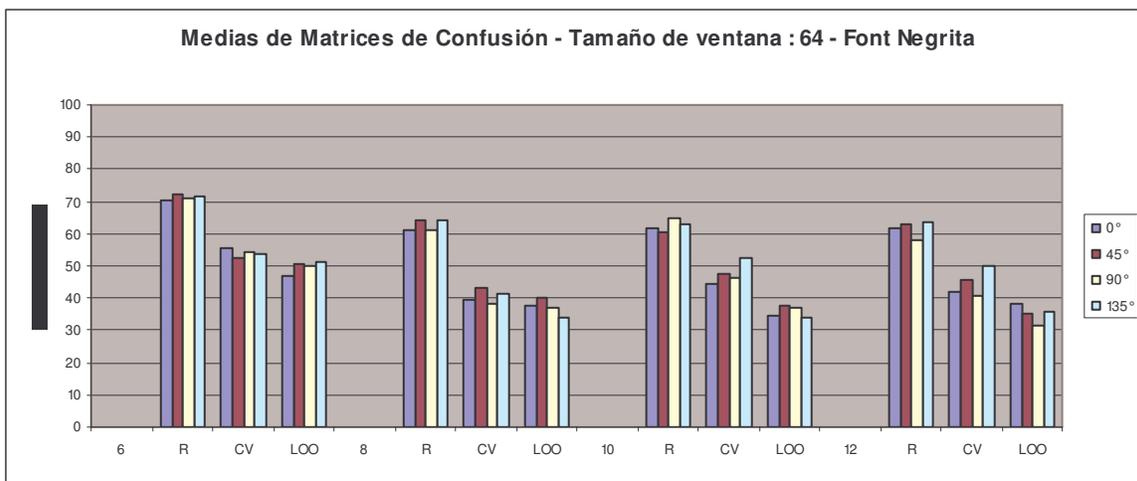
Tabla 4.36. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 256 píxeles, Font Itálica, Rotación 0°, 45°, 90°, 135°.



Grafica 4.26. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones 0°, 45°, 90° y 135°, Font Itálica

	0°	45°	90°	135°
6				
R	70.63	72.00	71.00	71.50
CV	55.75	52.25	54.50	54.00
LOO	47.13	50.38	50.25	51.25
8				
R	61.00	64.25	61.13	64.13
CV	39.25	43.25	38.25	41.25
LOO	37.50	40.25	36.88	33.88
10				
R	61.50	60.75	64.75	63.25
CV	44.75	47.75	46.50	52.50
LOO	34.38	37.63	37.00	33.75
12				
R	61.75	62.88	58.13	63.75
CV	42.00	45.50	41.00	50.00
LOO	38.00	35.13	31.63	36.00

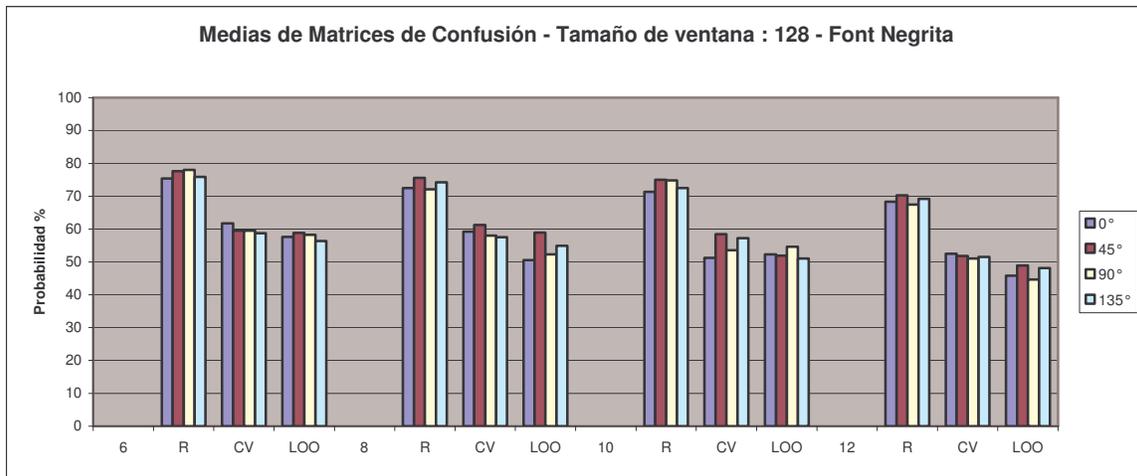
Tabla 4.37. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 64 píxeles, Font Negrita, Rotación 0°, 45°, 90°, 135°.



Grafica 4.27. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones 0°, 45°, 90° y 135°, Font Negrita

	0°	45°	90°	135°
6				
R	75.38	77.63	78.00	75.88
CV	61.75	59.50	59.50	58.75
LOO	57.63	58.88	58.25	56.38
8				
R	72.50	75.63	72.13	74.25
CV	59.25	61.25	58.00	57.50
LOO	50.50	59.00	52.25	54.88
10				
R	71.38	75.00	74.88	72.50
CV	51.25	58.50	53.50	57.25
LOO	52.25	51.88	54.63	51.00
12				
R	68.38	70.25	67.50	69.25
CV	52.50	51.75	51.00	51.50
LOO	45.75	48.88	44.63	48.13

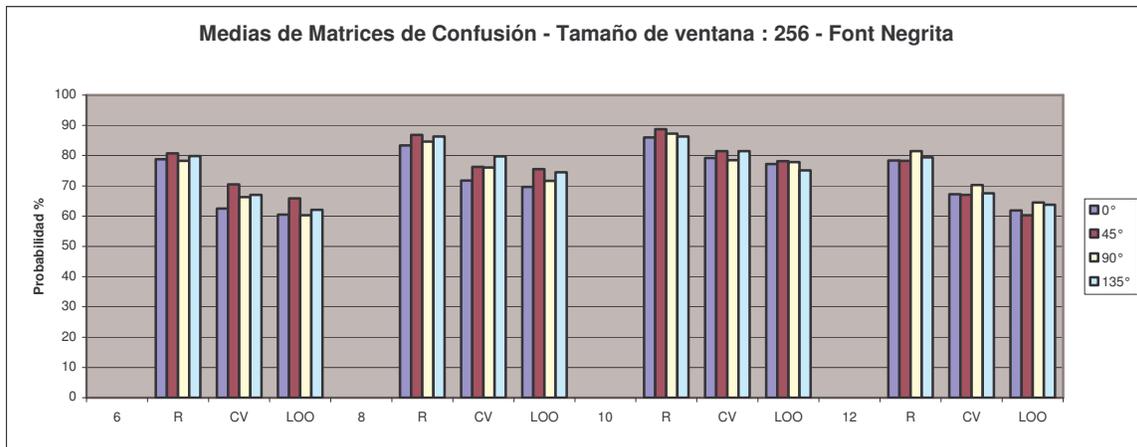
Tabla 4.38. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 128 píxeles, Font Negrita, Rotación 0°, 45°, 90°, 135°.



Grafica 4.28. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones 0°, 45°, 90° y 135°, Font Negrita

	0°	45°	90°	135°
6				
R	78.75	80.75	78.25	79.88
CV	62.50	70.50	66.25	67.00
LOO	60.50	65.88	60.25	62.13
8				
R	83.38	86.88	84.63	86.38
CV	71.75	76.25	76.00	79.75
LOO	69.63	75.50	71.63	74.50
10				
R	86.00	88.75	87.25	86.38
CV	79.25	81.50	78.50	81.50
LOO	77.25	78.13	77.88	75.13
12				
R	78.38	78.25	81.50	79.38
CV	67.25	67.00	70.25	67.50
LOO	61.88	60.25	64.50	63.75

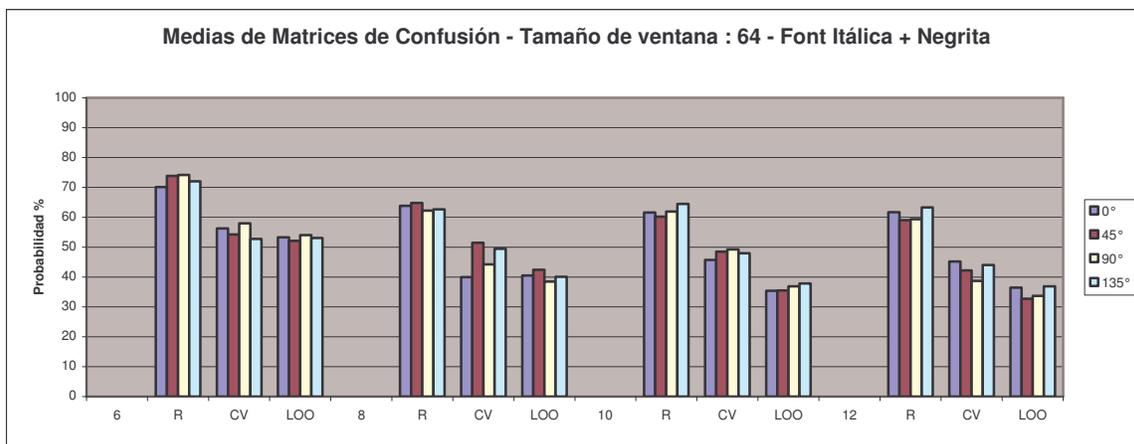
Tabla 4.39. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 256 píxeles, Font Negrita, Rotación 0°, 45°, 90°, 135°.



Grafica 4.29. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones 0°, 45°, 90° y 135°, Font Negrita

	0°	45°	90°	135°
6				
R	70.13	73.88	74.13	72.00
CV	56.25	54.25	58.00	52.75
LOO	53.25	52.13	54.00	53.13
8				
R	63.88	64.75	62.25	62.63
CV	40.00	51.50	44.25	49.50
LOO	40.50	42.50	38.50	40.13
10				
R	61.63	60.25	61.88	64.50
CV	45.75	48.50	49.25	48.00
LOO	35.38	35.50	36.88	37.88
12				
R	61.75	59.00	59.38	63.25
CV	45.25	42.25	38.75	44.00
LOO	36.50	32.75	33.75	36.88

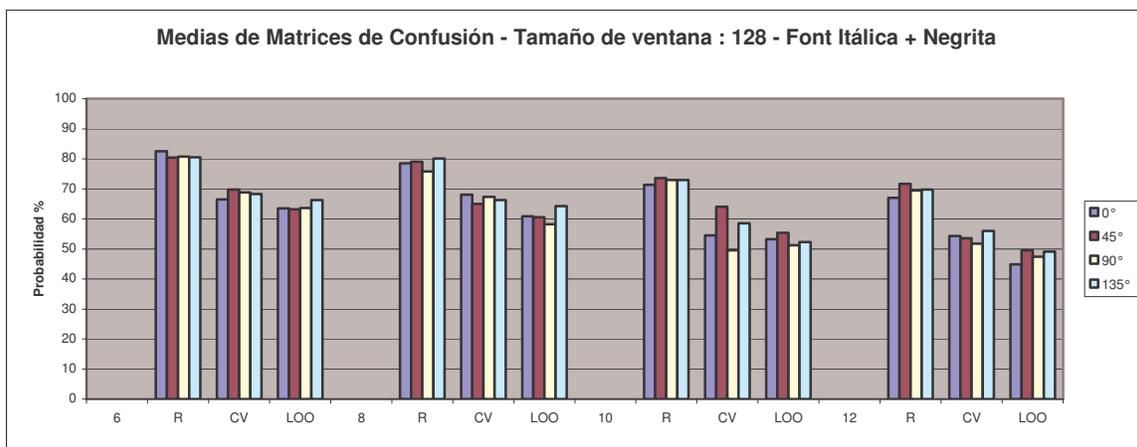
Tabla 4.40. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 64 píxeles, Font Itálica + Negrita, Rotación 0°, 45°, 90°, 135°.



Grafica 4.30. Medias de probabilidad de matrices de confusión, tamaño de ventana 64, con rotaciones 0°, 45°, 90° y 135°, Font Itálica + Negrita.

	0°	45°	90°	135°
6				
R	82.50	80.38	80.75	80.50
CV	66.50	69.75	68.75	68.25
LOO	63.50	63.13	63.63	66.25
8				
R	78.50	79.00	75.75	80.13
CV	68.00	65.00	67.25	66.25
LOO	60.88	60.50	58.25	64.25
10				
R	71.38	73.50	72.88	72.88
CV	54.50	64.00	49.50	58.50
LOO	53.25	55.38	51.25	52.25
12				
R	67.00	71.63	69.38	69.75
CV	54.25	53.50	51.75	56.00
LOO	44.88	49.50	47.38	49.13

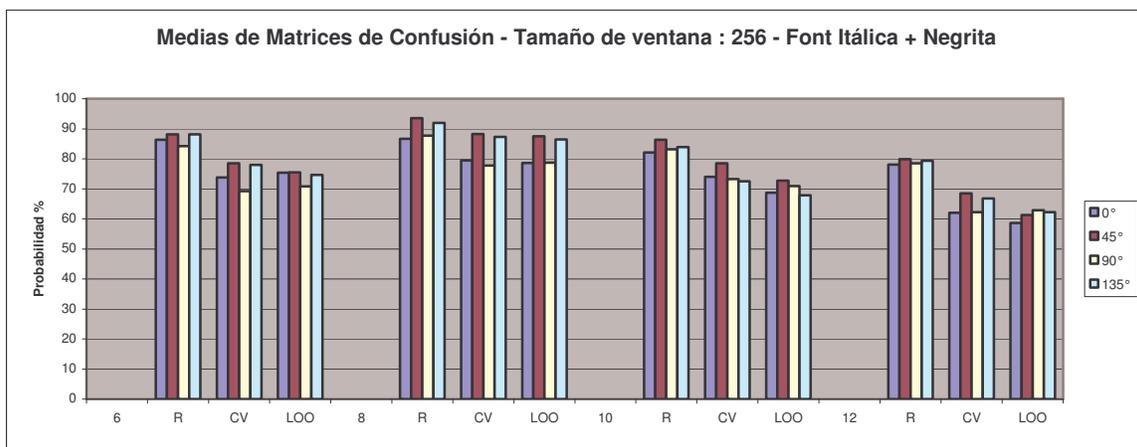
Tabla 4.41. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 128 píxeles, Font Itálica + Negrita, Rotación 0°, 45°, 90°, 135°.



Grafica 4.31. Medias de probabilidad de matrices de confusión, tamaño de ventana 128, con rotaciones 0°, 45°, 90° y 135°, Font Itálica + Negrita.

	0°	45°	90°	135°
6				
R	86.38	88.13	84.25	88.13
CV	73.75	78.50	69.25	78.00
LOO	75.38	75.50	70.75	74.63
8				
R	86.63	93.50	87.75	92.00
CV	79.50	88.25	77.75	87.25
LOO	78.63	87.50	78.75	86.50
10				
R	82.13	86.38	83.13	83.88
CV	74.00	78.50	73.25	72.50
LOO	68.63	72.75	70.88	67.88
12				
R	78.13	79.88	78.50	79.38
CV	62.00	68.50	62.25	66.75
LOO	58.63	61.25	62.88	62.25

Tabla 4.42. Tabla de medias probabilidades, clasificador modificado, tamaño de ventana de 256 píxeles, Font Itálica + Negrita, Rotación 0°, 45°, 90°, 135°.



Gráfica 4.32. Medias de probabilidad de matrices de confusión, tamaño de ventana 256, con rotaciones 0°, 45°, 90° y 135°, Font Itálica + Negrita.

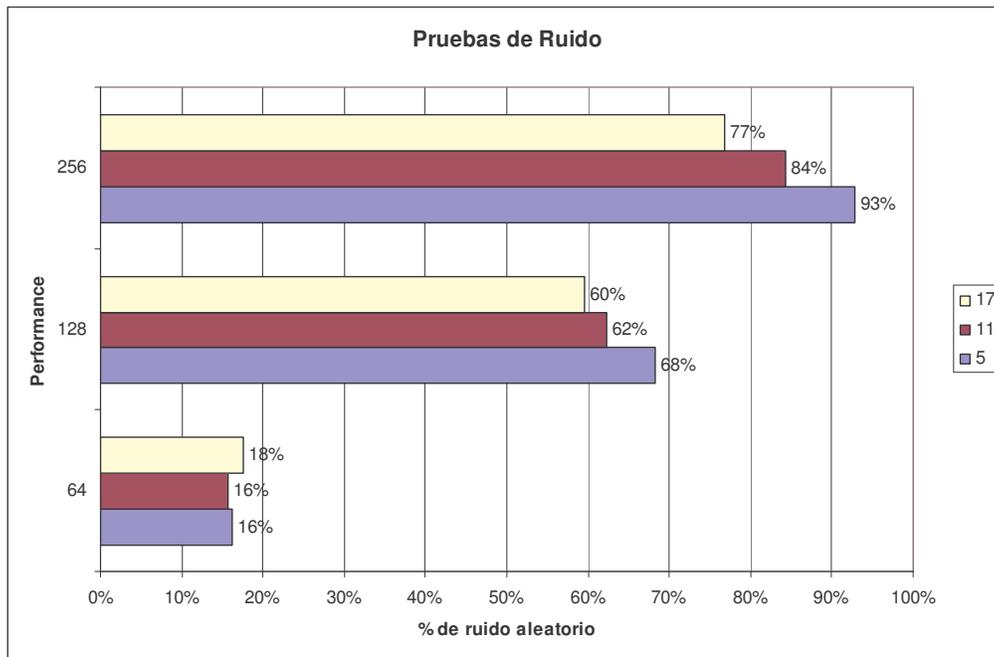
4.6. Resultados de sensibilidad al ruido.

Los resultados previos fueron obtenidos sin ruido aleatorio, pero es importante considerar el ruido ya que las imágenes son comúnmente contaminadas por procesos tales como la digitalización o scaneo, de baja calidad.

Realizamos un conjunto de pruebas sobre muestras con ruido aleatorio (al 5, 11 y 17%), en la Figura 4.19, aparecen unos ejemplos de ellas. Se hicieron 100 pruebas de ventanas aleatorias de dimensiones 256, 128 y 64 píxeles, sobre cada una de las imágenes con ruido. Los resultados se muestran en la Gráfica 4.33.



Figura 4.19. Muestras de textos con ruido aleatorio al 5% (a), 11% (b), 17% (c).



Gráfica 4.33. Resultados de pruebas de texto con ruido aleatorio al 5, 11 y 17%.

5. Conclusiones y perspectivas.

Llegamos satisfactoriamente a implementar la metodología propuesta, así también como a conformar el bloque de texto tras el preprocesamiento. Se implementaron satisfactoriamente los clasificadores Gaussiano y KNN, así como los métodos propuestos para la evaluación del performance.

Una vez obtenida la cadena del procesamiento, se pasó a probar los tipos de fuente elegidos, en los tamaños y orientaciones seleccionados, obteniendo los resultados de clasificación máximos para las ventanas de 256 píxeles, a 6 puntos en promedio, para todos los casos.

En promedio, la tasa de buena clasificación para el font regular e itálica a 0° de rotación quedo en 8 puntos a 300 DPI, y en 6 puntos con 0° de rotación para el caso de combinación de Negrita con Itálica, como una excepción.

No se obtuvo el nivel de performance que nos habíamos planteado anteriormente al inicio de este trabajo, un nivel superior al 95%, porque los momentos cambian, conforme a la rotación y a la escala, a pesar de que se llegaron a buenos resultados de clasificación.

Como perspectiva a futuro podemos pensar en pruebas con otro tipo de fonts y ver el performance en otros idiomas (Inglés, Alemán, Francés, etc.), probar a diferentes grados de rotación, ya que aquí lo hemos hecho a espacio de 45°, hacer pruebas a 5° o 10°, y estudiar más como afecta la variabilidad de los momentos. También se podría probar sobre otros juegos de caracteres, en otros idiomas (Chino, Arabe, etc.), como lo hacen otros autores [19], tratando de obtener una mejor identificación con descriptores de font para cada tipo de lenguaje.

Referencias Bibliográficas.

1. Aviles-Cruz Carlos, “Thèse: Analyse de Texture par Statistiques D’Ordre Supérieur: Caracterisation et Performances”, Institut National Polytechnique de Grenoble. 1997.
2. Aviles-Cruz Carlos, Rangel-Kuoppa Risto et al, “High-order statical texture analysis – Font recognition applied”, *Pattern Recognition Letters*, v.26, pp. 135-145. 2005.
3. Cooperman R. “Producing Good Font Attribute Determination Using Error-Prone Information”. In *Proceedings of the SPIE - Document Recognition IV*, pages 50--57, 1997.
4. Fakunaga Keinosuke, “Introduction to Statical Pattern Recognition”, Edit. Academic Press Inc. 1990.
5. González Rafael, Woods Richard, “Tratamiento Digital de Imágenes”, Edit. Addison-Wesley / Diaz de Santos. 1996.
6. Gueury Marc, Wijns Hervé, “Pretraitement et Classification de Donnees Application a la Discrimination de Textures”, *Mémoire de fin d’études*, Institut National Polytechnique de Grenoble. 1994
7. Hongwei Shi , Theo Pavlidis. “Font Recognition and Contextual Processing for More Accurate Text Recognition”, *Proceedings of the 4th International Conference on Document Analysis and Recognition*, p.39-44, August 18-20, 1997.
8. Julesz, B., Bergen, J.R., “Textons, The Fundamental Elements in Preattentive vision and Perception of Textures. *Bell Syst. Tech.J.* 62 (6), 1619-1645.
9. Khoubyari S. and Hull J. J. “Font and Function Word Identification in Document Recognition”. *Computer Vision and Image Understanding*, Vol. 63 no.1, pp. 66-74, 1996.
10. Malik, J. Belongie, S., Shi, J., Leung, T. “Textons, Contours and Regions: Cue Integration in Image Segmentation”. *IEEE Int. Conference on Computer Vision*, CORFO, Greece, September 1999.
11. Mathews John, “Métodos Numéricos con Matlab”, 3ra. Ed., Edit. Prentice Hall. 2000.
12. Mendenhall W., Wackerly D.D., Scheafer Richard., “Estadística Matemática con Aplicaciones”. Edit. Grupo Editorial Iberoamerica, 2da. Ed. 1994.
13. Nagy G., “Twenty Years of Document Image Analysis in PAMI”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, no. 1, pp. 38-62, Jan. 2000.
14. Nakamura S., “Análisis Numérico y Visualización Gráfica con Matlab”. Edit. Prentice Hall. 1997.
15. Schreyer, P. Suds, and C. Maderlechner, "Font Style Detection in Documents Using Textons", *Proc. Third Document Analysis Systems Workshop. Assoc. for Pattern Recognition Int'l*, 1998.
16. Song-Chun Zhu, “What Are Textons?”, *International Journal of Computer Vision*, Vol. 62, Issue 1-2. April-May 2005.
17. The Mathworks Inc., “Matlab Edición del Estudiante, Versión 4, guía de usuario”, Edit. Prentice may. 1996.
18. The Mathworks Inc., “Using Matlab, Ver. 6, Release 12, Reference Guide” . 2000.

19. Yong Zhu, "Font Recognition Based on Global Texture Analysis", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 23, no. 10, pp. 1192-1200, Oct. 2001.
20. Zramdini A., Ingold R., "Optical Font Recognition Using Typographical Features", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.20 n.8, p.877-882, August 1998.