

**UNIVERSIDAD AUTONOMA DE MADRID**

**ESCUELA POLITECNICA SUPERIOR**



**PROYECTO FIN DE CARRERA**

**CONSTRUCCIÓN DE UNA NUEVA BASE DE DATOS PARA  
EL RECONOCIMIENTO AUTOMÁTICO DE CARACTERES  
MANUSCRITOS Y GENERACIÓN DE RESULTADOS DE  
REFERENCIA**

**Sara García-Mina Martínez  
OCTUBRE 2014**



**CONSTRUCCIÓN DE UNA NUEVA BASE DE DATOS PARA  
EL RECONOCIMIENTO AUTOMÁTICO DE CARACTERES  
MANUSCRITOS Y GENERACIÓN DE RESULTADOS DE  
REFERENCIA**

**AUTOR: Sara García-Mina Martínez**

**TUTOR: Marta Gómez Barrero**

**PONENTE: Julián Fierrez Aguilar**

**Grupo ATVS**

**Dpto. Tecnología Electrónica y de las Comunicaciones**

**Escuela Politécnica Superior**

**Universidad Autónoma de Madrid**

**Octubre de 2014**



**Resumen:**

En este proyecto se adquiere una base de datos de caracteres manuscritos, complementaria a la base de datos MNIST, de libre acceso a la comunidad científica, y se evalúa el rendimiento de un sistema de referencia sobre la misma.

Tras una introducción al reconocimiento de caracteres manuscritos y un estudio del estado del arte del reconocimiento *offline* y bases de datos de referencia, se seleccionan las herramientas más representativas para llevar a cabo el proyecto.

Una vez adquirida la base de datos, se procesan las muestras usando herramientas de procesamiento de imágenes, para la posterior obtención de resultados de referencia con sistemas de reconocimiento de caracteres del estado del arte.

En la parte experimental se ha generado un conjunto de resultados para la base de datos adquirida. Dichos resultados servirán como referencia para evaluar el rendimiento de futuros sistemas de reconocimiento de caracteres manuscritos.

Se realizan diferentes experimentos comparando unos tipos de caracteres manuscritos u otros. Los diferentes tipos son:

- Números
- Mayúsculas
- Minúsculas

Para evaluar el rendimiento del sistema de referencia se han usado tasas de error, matrices de confusión y curvas CMC, que posteriormente serán explicadas, analizadas y comentadas. Para cerrar el apartado experimental, se realiza una comparación de resultados entre los distintos escenarios estudiados. Finalmente, se proponen conclusiones del trabajo realizado y trabajo futuro.

**Palabras clave:**

Base de datos, caracteres manuscritos, reconocimiento caracteres manuscritos, procesado de imágenes.



**Abstract:**

In this project a database is acquired, complementary to the MNIST database comprising handwritten characters of the Spanish alphabet. The database will be publicly available for the research community.

After an introduction to recognition of handwriting characters and a study of the state of the art of offline recognition and reference databases, the most representative tools are chosen to carry out the project.

The database is not only acquired but processed, using image processing tools to normalize it. The aim is to get a wide database and baseline results with recognition systems of handwritten characters of the state in the art.

In the experimental section a set of recognition result was generated for the database acquired. The results will be used as a reference to evaluate the performance of future algorithms on handwritten character recognition.

Different experiments were performed, depending on the different kinds of handwritten characters used for each experiment. The different kinds are:

- Numbers
- Capital letters
- Lowercase letters

The gathered results for the experimental section are the error rates of each experiment as well as the confusion matrix and the CMC Curves, which will be explained, analyzed and commented .To complete the experimental section a comparison between the different experiments is performed. Finally, conclusions and future work are exposed.

**Key words:**

Database, handwritten characters, handwritten character recognition, image processing.



## *Agradecimientos*

Quiero agradecer en primer lugar a mi profesor, Javier Ortega, la oportunidad que me ha brindado de colaborar con el ATVS desde 2010 como colaboradora en ámbitos de investigación y su posterior apoyo para realizar mi Proyecto Fin de Carrera en el laboratorio.

Me siento muy agradecida al resto de miembros del ATVS, por haberme permitido disfrutar como lo he hecho durante el desarrollo de este proyecto. En especial a Javier Galbally por apoyarme y animarme a lo largo de todo este tiempo mostrando interés por mi evolución con el proyecto fin de carrera.

Agradecer a mi tutora Marta Gómez Barrero, quien me ha prestado un apoyo y consejo determinantes y me han guiado de forma excepcional para la realización de este proyecto. También agradecer a mis compañeros de laboratorio en especial a mis compañeros y muy amigos Fátima García Donday, Ricardo Landriz y Sandra Uceda quienes me han acompañado y alegrado todo el tiempo en el laboratorio. A mi pareja de prácticas durante prácticamente toda la carrera María Lucena, realmente sin su ayuda y humor no habría sido posible. Un agradecimiento especial a mi mejor amigo de la universidad, Rodrigo López Manrique con el que he compartido muchos conocimientos y buenos ratos. Agradecer a mis amigos más especiales de la carrera Eva, Berta, Jaime, Helia, Pencho y Juanma, y al resto de compañeros.

Por último quería agradecer el apoyo mostrado por mi familia y amigos durante este proyecto y durante toda mi carrera universitaria. En especial a Jorge Borges que ha sido, destacadamente, mi mayor apoyo en el transcurso de la carrera y a mi amiga Victoria que me ha acompañado desde la infancia. A mis amigas y entrenadores del Olímpico Club de Rugby, en especial a Marta Iglesias, Jordan, Bilbo, Helen y Rico entre otras del club, me han apoyado y se han preocupado mucho por mí en todo momento.

Y para finalizar quiero dedicar este proyecto a mis padres Fermín y Almudena, que de verdad pienso que son los mejores del mundo y todo lo han dado todo por la educación y el progreso de sus hijos.



# INDICE DE CONTENIDOS

|   |     |
|---|-----|
| Índice de figuras .....   | x   |
| Índice de tablas .....  | xii |
| 1 Introducción.....   | 1   |
| 1.1 Motivación del proyecto .....   | 1   |
| 1.2 Objetivos y enfoque.....  | 2   |
| 1.3 Estructura de la memoria.....   | 3   |
| 2 Introducción al reconocimiento de caracteres manuscritos.....                   | 5   |
| 2.1 Reconocimiento offline de caracteres manuscritos .....                        | 5   |
| 2.1.1 Arquitectura de sistemas de reconocimiento de patrones.....                 | 6   |
| 2.1.2 Pre-procesado de imágenes.....  | 7   |
| 2.1.3 Reconocimiento de caracter.....   | 8   |
| 2.1.4 Clasificadores de caracteres.....   | 9   |
| 2.2 Bases de datos de caracteres manuscritos.....                                 | 9   |
| 2.2.1 MNIST.....  | 10  |
| 2.2.2 Firemaker.....  | 11  |
| 2.2.3 USPS.....   | 13  |
| 2.3 Evaluación del rendimiento de un sistema de reconocimiento de caracteres..... | 14  |
| 3 Adquisición y pre-procesado de la base de datos.....                            | 19  |
| 3.1 Adquisición.....  | 19  |
| 3.1.1 Herramienta Biógrafo .....  | 21  |
| 3.2 Pre-procesado .....   | 23  |
| 3.2.1 Binarización de las imágenes.....   | 24  |
| 3.2.2 Segmentación de caracteres .....  | 25  |
| 3.2.3 Normalización de las imágenes .....   | 27  |
| 4 Descripción de la base de datos.....  | 29  |
| 4.1 Base de datos adquirida con biógrafo.....                                     | 29  |
| 4.2 Base de datos pre-procesada.....  | 32  |
| 5 Resultados de referencia.....   | 35  |
| 5.1 Protocolo experimental.....   | 35  |
| 5.1.1 Protocolo Experimento 1 .....   | 37  |
| 5.1.2 Protocolo Experimento 2 .....   | 38  |
| 5.2 Sistema de reconocimiento de caracteres .....                                 | 39  |
| 5.3 Resultados de referencia .....  | 42  |
| 5.3.1 Resultados Experimento 1: indubitados vs indubitados .....                  | 43  |
| 5.3.2 Resultados Experimento 2: dubitados vs indubitados .....                    | 46  |
| 5.3.3 Discusión .....   | 48  |
| 6 Conclusiones y trabajo futuro.....  | 53  |
| Referencias .....   | 55  |



# Índice de figuras

---

- Figura 1-1: MNIST caracteres**
- Figura 2-1: Esquema de funcionamiento de un sistema de reconocimiento de patrones**
- Figura 2-2: Caracteres manuscritos**
- Figura 2-3: Caracteres manuscritos de la base de datos MNIST**
- Figura 2-4: Caracteres manuscritos del subcorpus A**
- Figura 2-5: Ejemplo de caracter del conjunto de test de la base de datos USPS mal segmentado**
- Figura 2-6: Imágenes de caracteres del conjunto de entrenamiento de la base de datos USPS**
- Figura 2-7: Imágenes de caracteres del conjunto de test de la base de datos USPS**
- Figura 0-8: Ejemplo de sistema de identificación**
- Figura 2-9: Ejemplo de ranking de scores**
- Figura 2-10: Ejemplo curva CMC**
- Figura 2-11: Ejemplo de matriz de confusión (Firemaker)**
- Figura 3-1: Fragmento de un documento manuscrito por un individuo indubitado en minúsculas**
- Figura 3-2: Fragmento de un documento manuscrito por un individuo indubitado en letras mayúsculas**
- Figura 3-3: Fragmento de un documento manuscrito por un individuo indubitado realizando operaciones matemáticas**
- Figura 3-4: Ejemplo de documento manuscrito de un individuo dubitado**
- Figura 3-5: Pantallazo de la herramienta de adquisición de muestras de Biógrafo**
- Figura 3-6: Ejemplo de imágenes procesadas de la base de datos MNIST**
- Figura 3-7: Ejemplo de imágenes sin procesar de la base de datos Biógrafo**
- Figura 3-8: Procesado de inicio a fin**
- Figura 3-9: Imagen de la base de datos Biógrafo (caracter manuscrito 5) y su correspondiente binarización con valor de umbral 128**
- Figura 3-10: Imagen de la base de datos Biógrafo (caracter manuscrito 5) y su correspondiente binarización con valor de umbral 200**
- Figura 3-11: Proceso de segmentación y conversión a escala de grises**
- Figura 3-12: Normalización paso a paso**
- Figura 4-1: Número de muestras por cada individuo indubitado. La recta roja indica el número medio: 257**
- Figura 4-2: Número de muestras por cada individuo dubitado. La recta roja indica el número medio: 191**
- Figura 4-3: Número de muestras por caracter (subcorpus indubitados). La recta roja indica el número medio: 2026**
- Figura 4-4: Número de muestras por caracter (subcorpus dubitados). La recta roja indica el número medio: 151**
- Figura 5-1: Conjuntos de entrenamiento y test. Protocolo experimental**
- Figura 5-2: Ejemplo en pequeña escala de clasificador PCA**
- Figura 5-3: Matriz de confusión del escenario Todo del Experimento 1 representada con colores**
- Figura 5-4: Curva CMC del escenario Todo del Experimento 1. En la posición 5 de la lista, hay una probabilidad de acierto del 95.18%**

**Figura 5-5: Matriz de confusión del escenario Todo del Experimento 2 representada con colores**

**Figura 5-6: Curva CMC del escenario Todo del Experimento 2. En la posición 5 de la lista, hay una probabilidad de acierto del 93.61%**

**Figura 5-7: Letras “S” mayúsculas de la base de datos procesada**

**Figura 5-8: Letras “s” minúsculas de la base de datos procesada**

# Índice de tablas

---

**Tabla 2-1: Tasas de error obtenidas con diferentes métodos utilizando MNIST**

**Tabla 2-2: Tasas de error obtenidas con diferentes métodos utilizando el subcorpus A**

**Tabla 2-3: Tasas de error obtenidas con diferentes métodos utilizando el subcorpus B**

**Tabla 2-4: Tasas de error obtenidas con diferentes métodos utilizando Base de datos USPS**

**Tabla 4-1: Cantidad de muestras por caracter numérico del conjunto indubitado**

**Tabla 4-2: Cantidad de muestras por caracter del alfabeto en mayúsculas del conjunto indubitado**

**Tabla 4-3: Cantidad de muestras por caracter del alfabeto en minúsculas del conjunto indubitado**

**Tabla 4-4: Cantidad de muestras por caracter numérico del conjunto dubitado**

**Tabla 4-5: Cantidad de muestras por caracter del alfabeto en mayúsculas del conjunto dubitado**

**Tabla 4-6: Cantidad de muestras por caracter del alfabeto en minúsculas del conjunto dubitado**

**Tabla 5-1: Cantidad de muestras por caracter de test para cada escenario del Experimento 1: indubitados vs indubitados**

**Tabla 5-2: Cantidad de muestras por caracter de test para cada escenario del Experimento 2: dubitados vs indubitados**

**Tabla 5-3: Valor de los parámetros para cada modalidad de evaluación**

**Tabla 5-4: Tasas de error de los 4 escenarios del Experimento 1**

**Tabla 5-5: Matriz de confusión del escenario Todo del Experimento 1**

**Tabla 5-6: Tasas de error de los 4 escenarios del Experimento 2**

**Tabla 5-7: Matriz de confusión del escenario Todo del Experimento 2**

**Tabla 5-8: Probabilidad de fallo según escenario**

**Tabla 5-9: Ejemplos de asignación de clases erróneas de caracteres del conjunto de test del escenario Todo**



# 1 Introducción

---

## 1.1 Motivación del proyecto

Hoy en día, la identificación de personas para el acceso a determinados servicios o instalaciones es una tarea muy importante. Además, con el creciente desarrollo de internet y de los dispositivos móviles, se hace más notable la necesidad de confirmar la identidad de una persona de manera remota. La ventaja principal de la autenticación biométrica es que lleva a cabo el reconocimiento a través de lo que uno “es” y no de lo que recuerda (p.ej un PIN) o posee (p.ej. una llave).

El estudio de la identificación de personas mediante rasgos biométricos, consiste en la asociación de una identidad a un individuo. Estos estudios e investigaciones están basados en sistemas de reconocimiento y autenticación biométrica. El funcionamiento de dichos sistemas, consiste en técnicas matemáticas y estadísticas sobre rasgos físicos o de conducta de un individuo, para verificar su identidad.

Los rasgos físicos o de conducta de un individuo se denominan rasgos biométricos y no pueden ser perdidos u olvidados por lo que son potentes herramientas en la identificación personal.

Entre todos los rasgos biométricos existentes, la escritura manuscrita es uno de los más utilizados en la actualidad ya que tiene una característica muy importante, la **aceptabilidad**.

Firmar o escribir en cualquier ambiente es generalmente aceptado por la sociedad, sin embargo otros rasgos biométricos pueden ser rechazados. No todo el mundo está dispuesto a dar su huella dactilar, especialmente en determinadas situaciones de la vida cotidiana, mientras que poca gente pone problemas a la hora de firmar o escribir para confirmar su identidad.

Por ello, el estudio de la identificación por escritura manuscrita es una de las líneas de investigación más extendidas dentro del campo de la biometría.

El desarrollo de esta línea de investigación conlleva dos tipos de consideraciones:

- **Identificación del escritor.** Identificación de la persona que ha escrito un determinado mensaje (¿quién lo ha escrito?).
- **Identificación de la escritura.** Identificación de caracteres del contenido de un determinado mensaje manuscrito (¿qué se ha escrito?).

Ambas están relacionadas ya que para la identificación de escritor es necesaria la investigación en campos de identificación de escritura. Por lo que cobra una gran importancia el estudio de los mecanismos de reconocimiento de caracteres.

Para la investigación y desarrollo de estos sistemas de identificación es necesario la utilización de grandes bases de datos que contengan una suficiente variabilidad entre muestras, representativa de las distintas formas de escribir de las personas. Además, el número de muestras debe ser suficiente para poder obtener resultados estadísticamente

significativos sobre el rendimiento de los sistemas de reconocimiento. Hay varias bases de datos disponibles públicamente para realizar análisis y experimentos y de esta manera mejorar los sistemas de identificación.

La base de datos MNIST [1] es de las más utilizadas para el desarrollo de sistemas de identificación de caracteres y por lo tanto, la mayoría están adaptados a las características propias de MNIST. Sin embargo, a pesar del gran valor de esta base de datos, tiene algunas limitaciones, como por ejemplo, que sólo contiene muestras de los 10 dígitos, y no del resto de caracteres del alfabeto occidental. Además, toda la base de datos fue adquirida sobre individuos de cultura angloparlante, con lo cual los resultados obtenidos sobre ella pueden tener ligeras desviaciones frente a muestras capturadas de personas con otra ascendencia.

Las imágenes normalizadas de la base de datos MNIST tienen el fondo negro y el caracter en blanco con medidas de 28x28 píxeles. [1]  
A continuación se puede ver un ejemplo:



**Figura 1-1: MNIST caracteres**

La motivación para este proyecto es principalmente la adquisición de una base de datos de caracteres manuscritos, partiendo de los documentos de los cuales se obtienen los caracteres hasta la organización y el procesado de la misma. Ya que esta base de datos será publicada, una de las mayores motivaciones es que se utilizará por un rango muy amplio de investigadores en el sector de reconocimiento de caracteres manuscritos, aportando a la comunidad científica un gran apoyo para el continuo desarrollo y mejora de los sistemas de reconocimiento. Por lo que las principales motivaciones son:

- Apoyar el desarrollo y mejora de la comunidad científica de la biometría.
- Realizar una base de datos más completa y heterogénea que MNIST.
- Aportar más variabilidad debido a que los resultados de MNIST están muy saturados.

## **1.2 Objetivos y enfoque**

El proyecto tiene como objetivo la adquisición y el pre-procesado de una nueva base de datos, complementaria a la base de datos MNIST, compuesta por caracteres manuscritos del alfabeto castellano que sirva como futura referencia para la evaluación y mejora de sistemas de reconocimiento automático de este tipo de muestras.

Esta base de datos, además de ser adquirida, es procesada utilizando herramientas de tratamiento de imagen para que tenga el mismo formato que MNIST. Dando lugar a una base de datos “cómoda” para su uso, ya que muchos investigadores, de la comunidad biométrica de caracteres manuscritos, están acostumbrados a utilizar sistemas de reconocimiento adaptados a las características de las imágenes de MNIST.

Por lo tanto se ha llevado a cabo un procesado de las imágenes de la base de datos adquirida para su adecuada normalización y uso en las herramientas correspondientes de reconocimiento.

Para la parte experimental se ha generado un conjunto de resultados de reconocimiento utilizando técnicas del estado del arte. Dichos resultados servirán como referencia para evaluar el rendimiento de futuros algoritmos.

Se han obtenido tasas de error del sistema de identificación tanto para números como letras mayúsculas y minúsculas de nuestra base de datos. Finalmente, se presentan las conclusiones y se proponen líneas de trabajo futuras.

### **1.3 Estructura de la memoria**

La memoria consta de los siguientes capítulos:

- El capítulo 1 es la introducción del proyecto, donde se explican las principales motivaciones del mismo y se introduce la materia del proyecto detallando los objetivos y el enfoque que se le ha dado.
- El capítulo 2 expone el estado del arte de sistemas basados en detección de caracteres manuscritos. Se detalla el proceso de reconocimiento *offline* ya que es el utilizado en el presente proyecto. Además, se presentan las características de tres bases de datos de caracteres manuscritos; dos de ellas referentes para la comunidad científica (MNIST y USPS) y la tercera una base de datos adquirida en una universidad alemana no tan referente como las anteriores (Firemaker), pero relevante para el proyecto ya que es de ámbito universitario. Para concluir este capítulo se explican los modos de evaluación de rendimiento de los sistemas de identificación.
- En el 3 capítulo se detalla la adquisición y el pre-procesado de la base de datos objetivo del presente proyecto. Se expone la herramienta utilizada para la adquisición así como los pasos para el procesado de las imágenes de la base de datos.
- El capítulo 4 muestra en detalle las características de la base de datos adquirida antes y después de ser procesada.
- En el 5 capítulo se exponen los resultados de referencia obtenidos para la base de datos adquirida, explicando el sistema de reconocimiento utilizado y el protocolo empleado para la obtención de dichos resultados.
- Finalmente en el capítulo 6 se muestran las conclusiones y el trabajo futuro que conlleva este proyecto.



## 2 Introducción al reconocimiento de caracteres manuscritos

---

El reconocimiento automático de escritura manuscrita ha atraído la atención de los investigadores desde la creación de los ordenadores. Hoy en día los avances tecnológicos así como los avances en investigación hacen posible el desarrollo de nuevos sistemas para reconocer caracteres escritos con resultados cada vez menores. [2]

Las conferencias más relevantes actualmente, a las que acude una parte importante de la comunidad de investigación, son:

- *International Conference on Frontiers in Handwriting Recognition (ICFHR)*
- *International Conference on Document Analysis and Recognition (ICDAR)*

Ambas aprobadas por el organismo IEEE (Instituto de Ingeniería Eléctrica y Electrónica). Estas conferencias, entre otras, hacen que los investigadores de detección de escritura se mantengan al día y mejoren los resultados.

El texto escrito puede ser detectado de manera *offline* sobre un trozo de papel, una imagen de un texto a través de escáner óptico, o sistemas inteligentes de reconocimiento de caracteres manuscritos. Otras características como el movimiento del bolígrafo y la presión que se aplica al escribir, pueden ser detectadas de manera *online* utilizando, por ejemplo, pantallas adaptadas para la escritura manuscrita o tablets.

Los sistemas de detección *offline* son menos precisos que los *online*. Sin embargo, tienen un gran impacto económico en la actualidad para dominios especializados, como para interpretar las direcciones de texto manuscrito que hay en los sobres del sistema tradicional de correos, reconocimiento de texto en cheques bancarios, matriculas de coche, etc. [3]

En este apartado nos centraremos en detección *offline* ya que es la metodología que se ha usado en el desarrollo del proyecto.

### 2.1 Reconocimiento offline de caracteres manuscritos

Para comprender los sistemas de reconocimiento de caracteres manuscritos es necesario entender la arquitectura general de un sistema de reconocimiento de patrones y los modos de funcionamiento del mismo.

Así, tras describir el funcionamiento general de un sistema de reconocimiento de patrones, pasamos a detallar las funciones principales del reconocimiento offline. El reconocimiento de caracteres de un texto manuscrito es una tarea sencilla cuando el documento del que partimos está en buen estado y la letra es clara, pero esto no es siempre así.

El primer paso necesario para el reconocimiento es el problema de localizar el carácter del texto. Si el espacio de escritura no es muy amplio, hay mucho ruido en las imágenes o cerca de caracteres hay trazos espurios, es necesario un pre-procesado de las imágenes antes de pasar al reconocimiento de las mismas. Una vez procesadas las imágenes se pasa al reconocimiento de los caracteres extrayendo las características del mismo para crear un

patrón. Finalmente se procede a la tarea de clasificación, donde se les asigna el modelo más similar encontrado por el sistema.

### 2.1.1 Arquitectura de sistemas de reconocimiento de patrones

Antes de entrar en detalle en la arquitectura general de sistemas de reconocimiento de patrones, cabe destacar que dichos sistemas tienen diferentes modos de funcionamiento.

#### 2.1.1.1 Modos de funcionamiento

Dependiendo del contexto de aplicación un sistema de reconocimiento de patrones puede operar en modo registro o en modo identificación:

- **Registro:** Consiste en almacenar los patrones de las imágenes en el sistema. Para ello es necesario tener una base de datos con las imágenes de muestras requeridas para el sistema en cuestión. Este conjunto de imágenes se denomina conjunto de entrenamiento, de tal forma que cuanto mayor sea este conjunto de entrenamiento más robusto es el sistema de reconocimiento, más información tiene de cada muestra.
- **Identificación:** las imágenes a identificar se comparan con las del conjunto de patrones ya guardados, este proceso se conoce también como uno-a-muchos (1:N). Sin conocer la identidad presunta de la muestra, se compara uno a uno con los patrones ya existentes en el banco de datos registrados. El resultado de este proceso es la identidad de la muestra. En situaciones como identificación de escritura, el resultado del reconocimiento es averiguar el valor alfanumérico de cada carácter.

#### 2.1.1.2 Arquitectura

La arquitectura general de un sistema de reconocimiento de patrones comienza con la adquisición y el pre-procesado de las muestras. A continuación se extraen las características creando un patrón o modelo. Una vez obtenido el modelo, se compara los patrones almacenados en la base de datos (figura 2-1).

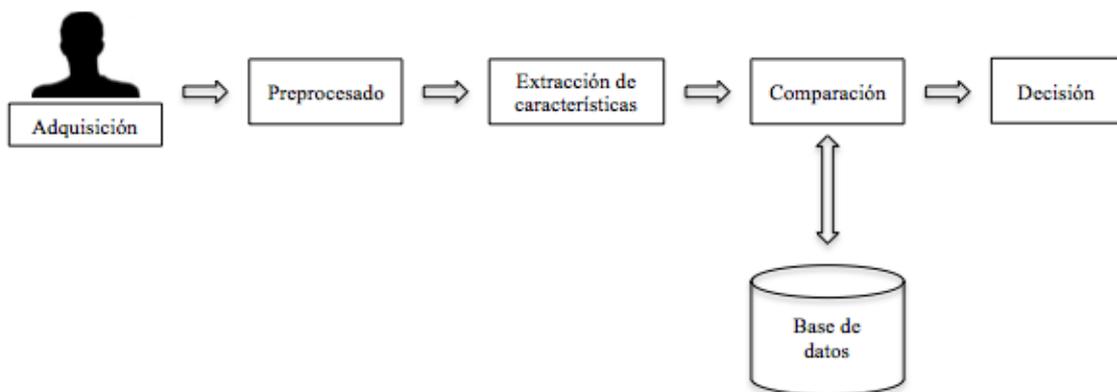


Figura 2-1: Esquema de funcionamiento de un sistema de reconocimiento de patrones.

- **Adquisición:** mediante instrumentos de adquisición, como pueden ser sensores, fotografías, vídeo, etc. se captura la muestra en cuestión para darle un formato digital. Este proceso es importante ya que la cantidad y calidad de las capturas será determinante en el rendimiento de los siguientes procesos así como en el rendimiento final del sistema de reconocimiento de patrones.
- **Pre-procesado:** consiste en un conjunto de operaciones realizadas a la señal digital obtenida para ajustarla a los siguientes procesos del sistema, normalizándola a los requisitos del sistema y adaptándola para su mejor identificación (binarizar, eliminar ruido, aumentar contraste, etc.).
- **Extracción de características:** una vez procesas las señales digitales se aplican algoritmos de extracción de características para obtener los patrones, o modelos, determinantes de cada señal. Estos patrones se agrupan en vectores, por lo que cada señal digital pasa a ser un vector de características que la definen. El resto de información de la señal digital es desechada.
- **Comparación de patrones:** en esta etapa se comparan los vectores de características, o modelos, obtenidos en la fase previa con los vectores de características almacenados para determinar coincidencias. Se compara el vector de características con todos los registrados en la base de datos del sistema obteniendo valores de similitud con cada uno de ellos.
- **Decisión:** una vez obtenidos los valores de similitud, denominados *scores*, hay que tomar una decisión. Se comparan todos los *scores* obtenidos del modelo introducido, con respecto a cada patrón posible, mostrando una lista jerárquica con los patrones de mayor a menor valor de similitud, identificando la muestra introducida con el modelo que ha obtenido mayor *score* (el primer patrón de la lista jerárquica).

### 2.1.2 Pre-procesado de imágenes

Una vez comprendida la arquitectura general de un sistema de reconocimiento de patrones, pasamos a explicar las distintas tareas a realizar en el reconocimiento *offline* de caracteres manuscritos, comenzando con el pre-procesado de las imágenes. Es necesario realizar varias operaciones de análisis de documentos antes de reconocer el texto escaneado. Algunas operaciones comunes antes del reconocimiento de texto son [4]:

- **Umbralización:** es la tarea de convertir una imagen en escala de grises en una imagen binaria en blanco y negro. Consiste en diferenciar el carácter del fondo de la imagen. El histograma de la escala de grises de un documento, normalmente consiste en dos picos: uno que corresponde al fondo de la imagen y otro que corresponde al primer plano, a la imagen de interés. La tarea de obtener el valor de umbral que binariza la escala de grises (por encima del umbral, todos los valores de la escala de grises se convierten en blanco y por debajo del umbral, todos los valores de la escala de grises se convierten en negro) consiste en determinar un valor óptimo que se encuentra entre estos dos picos del histograma.
- **Eliminación de ruido:** la captura digital de imágenes puede introducir ruido debido a los dispositivos de escáner y la transmisión de las imágenes. Eliminar el ruido es la tarea de extraer la imagen principal desechando lo que no corresponde al texto. Consiste en eliminar la textura del fondo de la imagen. Esta textura se conoce como ruido y su aspecto se asemeja a imágenes de sal y pimienta. Puede haber ruido por el fondo de la imagen o interfiriendo el carácter [5].

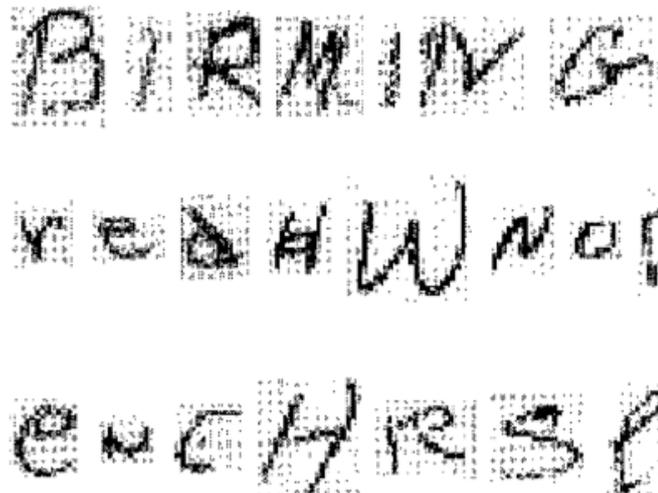
- **Segmentación de caracteres:** la segmentación de texto manuscrito en caracteres tiene varios enfoques. Numerosos métodos de detección de caracteres digitales consisten en identificar espacios entre caracteres y palabras. Sin embargo, en el dominio de texto manuscrito puede complicarse el proceso de segmentación debido a que los trazos de caracteres cercanos pueden intersectarse con las líneas del deseado. Un método muy usado en la actualidad son los filtros de Gabor, utilizados en aplicaciones como segmentación de caracteres alfanuméricos de matrículas de vehículos [6].

### 2.1.3 Reconocimiento de caracter

Una vez pre-procesadas las imágenes se procede al reconocimiento de carácter. El problema básico es asignar el caracter digitalizado a su clase simbólica. En el caso de una imagen impresa este problema se conoce como reconocimiento óptico de caracteres (OCR). En el caso de escritura manuscrita, se denomina reconocimiento inteligente de caracteres (ICR).

Los caracteres se clasifican tradicionalmente en mayúsculas, minúsculas, números y símbolos especiales como la exclamación, los paréntesis, etc. Se utilizan algoritmos de reconocimiento de patrones para extraer las características del caracter y asignarle la clase a la que pertenece. Las redes neuronales artificiales han emergido como métodos rápidos para implementar clasificadores para sistemas OCR [7]. Los algoritmos basados en el método “el vecino más cercano” son más precisos pero más lentos.

El reconocimiento de un caracter de máquina puede hacerse con mucha exactitud. La dificultad llega cuando se tienen que reconocer caracteres manuscritos. Algunos ejemplos de caracteres manuscritos se muestran en la figura 2-2



**Figura 2-2: Ejemplo de caracteres manuscritos**

El rendimiento del reconocimiento de caracter depende del enfoque de extracción de características, siendo éstos muy numerosos [8].

Varios experimentos muestran que la dirección del trazo es una de las características más eficientes para el reconocimiento de caracteres manuscritos [9]. Esta característica es

básicamente estadística y se representa con histogramas de elementos locales direccionales.

Otras mejoras se logran mediante la característica de curvatura [10] o estructura local [11-13]. Las propiedades locales y estructurales de un carácter son almacenadas en un vector de características, que corresponde a un punto en el espacio de características.

#### **2.1.4 Clasificadores de caracteres**

Por último, la tarea de clasificar consiste en particionar el espacio de características en regiones correspondientes a las distintas clases de caracteres. Las técnicas estadísticas [14-16] y las redes neuronales [17,18] han sido muy utilizadas para la tarea de clasificación debido a la eficiencia de la implementación. Los clasificadores estadísticos están divididos en paramétricos y no paramétricos. Incluyen la función lineal discriminante (LDF *Linear Discriminator Function*), la función cuadrática discriminante (QDF *Quadratic Discriminant Function*), el método de clasificación vecino más cercano (1-NN *Nearest-Neighbor*), etc.

La función modificada cuadrática discriminante (MQDF2) propuesta por Kimura ha sido muy usada debido a que reduce la complejidad y mejora el rendimiento de la clasificación de la función QDF [19,20].

La clasificación con redes neuronales incluyen perceptrón multicapa (MLP *Multilayer Perceptron*) [21], la función base radial (RBF *Radial Basis Function*) [17], el clasificador polinómico [22] etc.

La principal diferencia entre clasificadores estáticos y clasificadores de redes neuronales es que los parámetros de los clasificadores de redes neuronales están optimizados en el aprendizaje supervisado discriminativo con el objetivo de separar los patrones de diferentes clases. Si la estructura de la red se ha diseñado correctamente y el número de muestras de entrenamiento es suficientemente grande, las redes neuronales son capaces de clasificar los caracteres de test de una forma muy precisa.

En los últimos años un nuevo tipo de clasificador, llamado *Support Vector Machine* (SVM) [23], está siendo muy utilizado. Un clasificador SVM se basa en la teoría de aprendizaje estadística de Vapnik [24] y la optimización de la programación cuadrática. SVM es básicamente un clasificador binario. Múltiples SVMs se pueden combinar para dar lugar a un sistema de clasificación de múltiples clases.

## **2.2 Bases de datos de caracteres manuscritos**

Existen multitud de bases de datos de caracteres manuscritos de distintos idiomas y culturas. Cada base de datos tiene sus características específicas que la hacen más o menos apropiadas para determinados usos.

La variedad de bases de datos es muy útil, ya que se pueden utilizar para mejorar los algoritmos de clasificación y reconocimiento. Cuanta más variedad, más robusto y eficaz podemos hacer un programa de reconocimiento.

En la actualidad la base de datos más conocida y utilizada es MNIST [1], una base de datos estadounidense de caracteres manuscritos numéricos que se detallará a continuación.

En los siguientes apartados se exponen las características de tres bases de datos diferentes: MNIST, USPS muy utilizadas mundialmente y Firemaker, una base datos alemana no tan conocida como las dos anteriores, pero usada en el ámbito académico.

### 2.2.1 MNIST

La base de datos de caracteres numéricos manuscritos MNIST es pública [1]. Está compuesta por un conjunto de 60.000 imágenes de caracteres de entrenamiento y 10.000 imágenes de caracteres de test.

Las imágenes originales en blanco y negro obtenidas de NIST se normalizaron a un tamaño de 20x20 píxeles. Estas fueron procesadas para convertirlas a escala de grises, con un algoritmo de normalización que utilizaba una técnica “anti-aliasing”. Las imágenes fueron centradas en un marco de 28x28 utilizando la técnica de centro de masas.

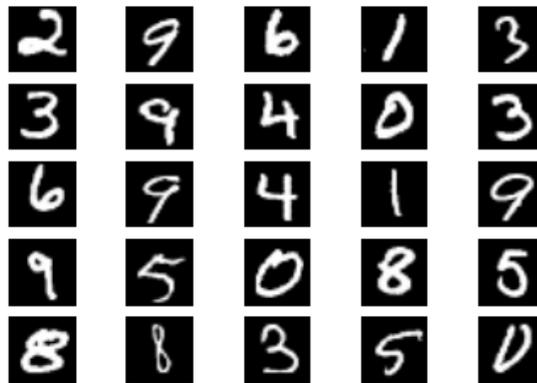


Figura 2-3: Caracteres manuscritos de la base de datos MNIST

MNIST fue construida a partir de la *Special Database 1* (SD-1) y 3 (SD-3) de NIST, que contienen imágenes binarias de caracteres manuscritos. Originalmente se diseñó SD-3 como un conjunto de entrenamiento y SD-1 como un conjunto de test. Sin embargo, SD-3 contiene imágenes con caracteres más precisos y fáciles de reconocer que SD-1. Esto es debido a que SD-3 fue obtenida de empleados de la empresa Census Bureau, mientras que la SD-1 fue obtenida mediante estudiantes de instituto.

Por ello, MNIST mezcló ambas bases de datos en sus conjuntos de entrenamiento y test. El conjunto de entrenamiento de MNIST está compuesto por 30.000 imágenes de caracteres SD-3 y 30.000 caracteres de SD-1. El de test, por 5.000 imágenes SD-3 y 5.000 imágenes de SD-1. Las 60.000 imágenes de caracteres del conjunto de entrenamiento contienen ejemplos de 250 escritores. Los escritores del conjunto de test son distintos a los de entrenamiento.

Multitud de métodos han sido testeados con la base de datos MNIST ya que es el estándar referencial para los algoritmos reconocedores y clasificadores de caracteres manuscritos. En la tabla 2-1 se muestran algunos de los resultados. [1]

| <b>Clasificador</b>                              | <b>Tasa de error (%)</b> |
|--|--------------------------|
| Clasificador lineal (1-NN)                       | 12.0                     |
| K-Nearest-Neighbors (LeCun et al.)               | 5.0                      |
| K-Nearest-Neighbors (Kenneth Wilder)             | 2.83                     |
| PCS+ Quadatric classifier                        | 3.3                      |
| SVM, Gaussian Hernel                             | 1.4                      |
| SVM (DeCoste and Scholkopf)                      | 0.68                     |
| Redes Neuronales (LeCun et al. 1998)             | 2.45                     |
| Redes Neuronales (Simard et al. ICDAR 2003)      | 1.1                      |
| Redes Neuronales (Deng et al. Interspeech 2010)  | 0.83                     |
| Redes Convolucionales (LeCunt et al.1998)        | 1.1                      |
| Redes Convolucionales (Simard et al. ICDAR 2003) | 0.4                      |

**Tabla 2-1: Tasas de error obtenidas con diferentes métodos utilizando MNIST**

## 2.2.2 Firemaker

La base de datos Firemarker consiste en caracteres manuscritos de 251 estudiantes alemanes. Cada estudiante escribió cuatro páginas diferentes:

1. Texto manuscrito específico.
2. Texto manuscrito específico en mayúsculas.
3. Texto manuscrito específico falsificando el estilo de escritura.
4. Texto manuscrito libre.

Para la composición de la base de datos, solo fueron usadas las segundas páginas.

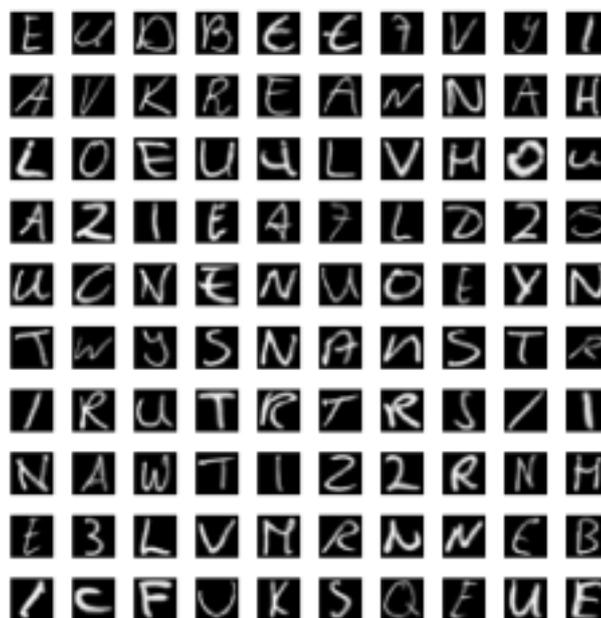
Las páginas manuscritas contienen un texto en alemán con 612 caracteres alfanuméricos, las letras son en mayúsculas. Los 251 estudiantes escribieron el mismo texto, este texto fue prediseñado para que contuviese gran variedad de caracteres alfanuméricos. Todos los estudiantes utilizaron el mismo tipo de papel y bolígrafo. Estos textos fueron escaneados en escala de grises.

Los caracteres están centralizados y normalizados. La normalización de los caracteres está hecha de dos maneras:

- Normalización centralizando los caracteres en el centro la imagen: el caracter se sitúa en una imagen cuadrada de tal forma que el número de pixeles por arriba y por abajo del caracteres son iguales, y el número de pixeles por la izquierda y por la derecha del caracter son iguales, a todos los lados del caracteres se le añaden 3 píxeles dando lugar a la imagen final de 56x56 pixeles. Las imágenes resultantes de esta normalización están almacenadas en el subcorpus A.
- Normalización centralizando los caracteres mediante el método centro de masas: el caracter se sitúa en una imagen cuadrada de tal forma que el centro de masas del objeto está localizado en el centro de la misma. La imagen cuadrada donde los caracteres están centrados tienen un tamaño de 90X90 pixeles. Las imágenes resultantes de esta normalización están almacenadas en el subcorpus B.

De esta manera existen dos subcorpus en la base de datos Firemaker, A y B.

Firemaker, en su conjunto, contiene 40.121 caracteres manuscritos. El caracter X no está incluido en esta base de datos, ya que es un carácter no utilizado en alemán. En la figura 2-4 se pueden observar algunos ejemplos de imágenes de caracteres del subcorpus A [25].



**Figura 2-4: Caracteres manuscritos del subcorpus A**

En las tablas 2-2 y 2-3 se presentan algunos resultados de experimentos en los dos subcorpus. Para la obtención de estos resultados, se utilizaron las bases de datos variando el conjunto de caracteres de entrenamiento y el conjunto de caracteres de test en un total de diez iteraciones para cada clasificador (validación cruzada con  $k=10$ ). Se dividió cada base de datos en diez conjuntos y en cada iteración, el conjunto de entrenamiento variaba. De tal manera que para cada clasificador se obtuvieron diez tasas de error.

En las tablas 2-2 y 2-3 se muestra el error generalizado, que consiste en la media de las diez tasas de error que se obtienen para cada clasificador y la varianza. [25]

| <b>Clasificador</b>                      | <b>Tasa de error generalizada (%)</b> |
|--|---------------------------------------|
| 1-NN (vecino más cercano)                | $21.68 \pm 0.61$                      |
| 3-NN (vecino más cercano)                | $20.79 \pm 0.44$                      |
| 5-NN (vecino más cercano)                | $20.74 \pm 0.64$                      |
| LKC ( <i>Linear Kerner Classifiers</i> ) | $32.99 \pm 0.74$                      |

**Tabla 2-2: Tasas de error obtenidas con diferentes métodos utilizando el subcorpus A**

| <b>Clasificador</b>                      | <b>Tasa de error generalizada (%)</b> |
|--|---------------------------------------|
| 1-NN (vecino más cercano)                | $17.79 \pm 0.53$                      |
| 3-NN (vecino más cercano)                | $17.23 \pm 0.82$                      |
| 5-NN (vecino más cercano)                | $17.28 \pm 0.66$                      |
| LKC ( <i>Linear Kerner Classifiers</i> ) | $30.20 \pm 0.55$                      |

**Tabla 2-3: Tasas de error obtenidas con diferentes métodos utilizando el subcorpus B**

### 2.2.3 USPS

USPS es una base de datos de caracteres numéricos obtenida tras escanear dígitos manuscritos de empleados del servicio de correos de los Estados Unidos. Las imágenes escaneadas originales son binarias, de diferentes tamaños y orientaciones del caracter.

El proceso de segmentación realizado por el servicio postal causó que algunas imágenes de dígitos no fuesen segmentadas. Debido a esto, la base de datos fue generada con dificultad para reconocer y clasificar los caracteres. En 1993, Simard et al. calculó una tasa de error humana de un 2.5%. Más adelante, se procesaron y normalizaron las imágenes, dando lugar a imágenes de tamaño 16x16 en escala de grises [26].

Hay 7291 caracteres de entrenamiento y 2007 caracteres de test. Estos conjuntos fueron adquiridos de forma ligeramente distinta. El conjunto de entrenamiento contiene muestras más precisas y fáciles de reconocer que el de test, en el que se pueden observar imágenes con caracteres no segmentados correctamente como se observa en la Figura 2-5



**Figura 2-5: Ejemplo de caracter del conjunto de test de la base de datos USPS mal segmentado.**

En las Figuras 2-5 y 2-6 se muestran imágenes de caracteres del conjunto de entrenamiento y el conjunto de test respectivamente. Se puede apreciar que el conjunto de test es mucho menos claro que el conjunto de entrenamiento, algunas imágenes del conjunto de test no se pueden ni identificar visualmente.



**Figura 2-6: Imágenes de caracteres del conjunto de entrenamiento de la base de datos USPS**



**Figura 2-7: Imágenes de caracteres del conjunto de test de la base de datos USPS**

En la tabla 2-4, se muestran algunos resultados experimentales realizados con la base de datos USPS [26].

| <b>Clasificador</b>            | <b>Tasa de error (%)</b> |
|--------------------------------|--------------------------|
| Redes Neuronales               | 4.2                      |
| Distancia tangencial           | 2.5                      |
| Vecino más cercano (NN)        | 5.6                      |
| K-NN, representaciones locales | 3.0                      |
| Kernel+Distancia tangencial    | 3.0                      |
| Lineal                         | 7.2                      |

**Tabla 2-4: Tasas de error obtenidas con diferentes métodos utilizando Base de datos USPS**

### **2.3 Evaluación del rendimiento de un sistema de reconocimiento de caracteres**

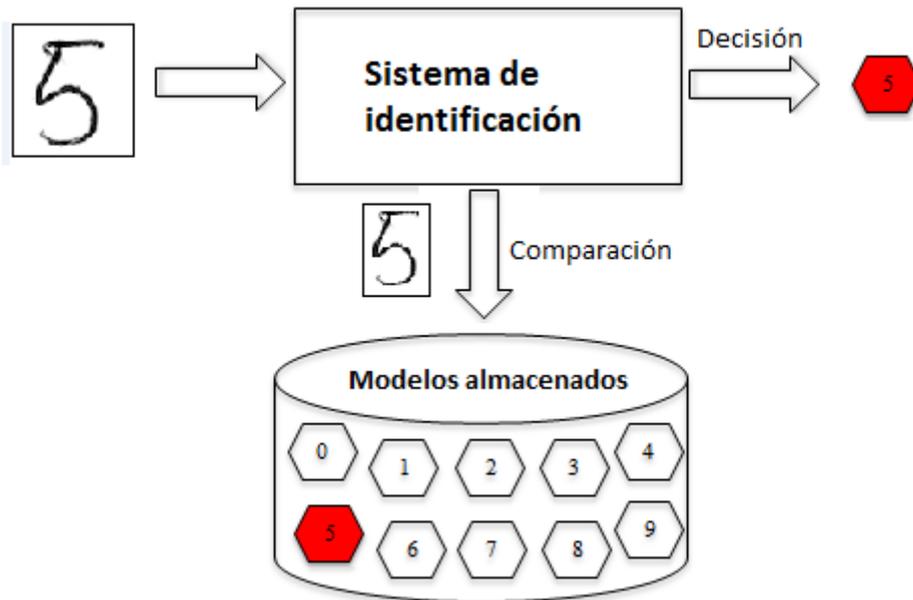
El rendimiento de un sistema de identificación puede verse afectado por los errores de detección que se pueden producir en el sistema debido a que las muestras pueden variar dando lugar a equivocaciones en la decisión del sistema. Las causas de esta varianza pueden ser diversas, como por ejemplo, la suciedad, el estado de ánimo, el instrumento de captación de la muestra, etc. Por lo que no siempre se obtiene el resultado esperado.

En sistemas de identificación se obtienen tantos valores de similitud, *scores*, como modelos haya en la base de datos. Tras la etapa de comparación de patrones o modelos, se obtienen los *scores* de la muestra introducida con respecto a cada modelo almacenado en el sistema. Se analizan los *scores* obtenidos seleccionando el valor de similitud más alto y se comunica la identificación obtenida.

Los resultados que dan los sistemas de identificación no son siempre óptimos; ya que una muestra puede variar en cada toma hasta el punto de parecerse más a otro modelo que a su correspondiente almacenado en la base de datos. Como ejemplo, si hablamos de caracteres manuscritos, una “i” puede llegar a identificarse como una “j” fácilmente, ya que la gráfica es parecida dependiendo de la forma de escribir que tenga cada persona.

En la figura 2-8 se introduce un “5” manuscrito, el sistema lo compara con todos los modelos que tiene almacenados en la base de datos y decide por el modelo que tiene un valor de similitud (*score*) mayor con la muestra introducida. Este caso es un ejemplo de un sistema de identificación funcionando correctamente.

Los sistemas de identificación obtienen un ranking de *scores* de la muestra introducida frente a todos los posibles modelos almacenados en la base de datos. De esta forma crean una lista jerárquica, o ranking de *scores*, con los modelos posibles de mayor a menor parecido con la muestra introducida y seleccionan el primer modelo de la lista, el que tiene mayor *score*, como modelo reconocido.



**Figura 2-8: Ejemplo de sistema de identificación.**

Para medir la cantidad de fallos de un sistema de identificación se utiliza el ER (*Error Rate*) también conocido como tasa de error. Este valor se obtiene realizando el cociente del número de fallos cometidos en la decisión, respecto al número total de pruebas realizadas. Si tenemos un sistema al que le introducimos 10 muestras y solo identifica correctamente 7 de ellas, el valor ER es  $3/10$ , 0.3. También se suele mostrar en tanto por ciento por lo que en este caso, el ER del sistema sería 30%.

Otra manera para medir el rendimiento del sistema de identificación es la utilización de curvas CMC (*Cumulative Match Curves*). Con la tasa de error sólo se evalúa la probabilidad de que el sistema acierte en la primera posición del ranking de *scores*. Sin embargo, con las curvas CMC se puede observar la probabilidad de que el modelo acertado este entre las “x” primeras posiciones de la lista jerárquica.

Las curvas CMC se obtienen a partir de varias identificaciones de distintas muestras, almacenando en qué posición del ranking de *scores* se encuentra el modelo acertado para cada muestra introducida en el sistema. En el ejemplo de la figura 2-9 se puede observar el ranking de *scores* obtenido para 4 muestras que se han introducido en un sistema de identificación de ejemplo. Sólo dos de estas muestras han sido acertadas en la primera posición, otra de ellas en la segunda y otra en la cuarta. De tal manera que la curva CMC para este ejemplo se presenta en la figura 2-10, donde se puede observar la probabilidad de que la muestra acertada esté entre las “x” primeras posiciones.

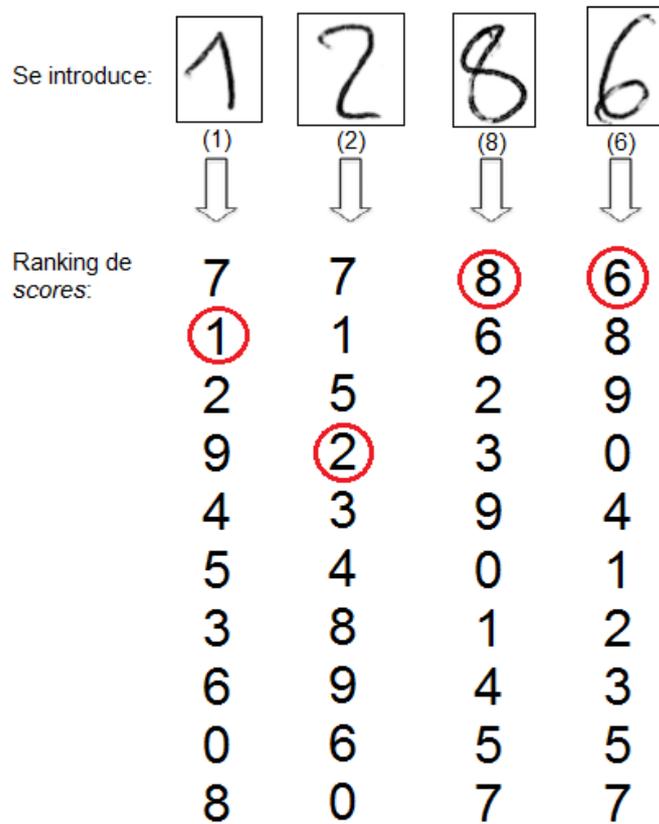


Figura 2-9: Ejemplo de ranking de *scores*

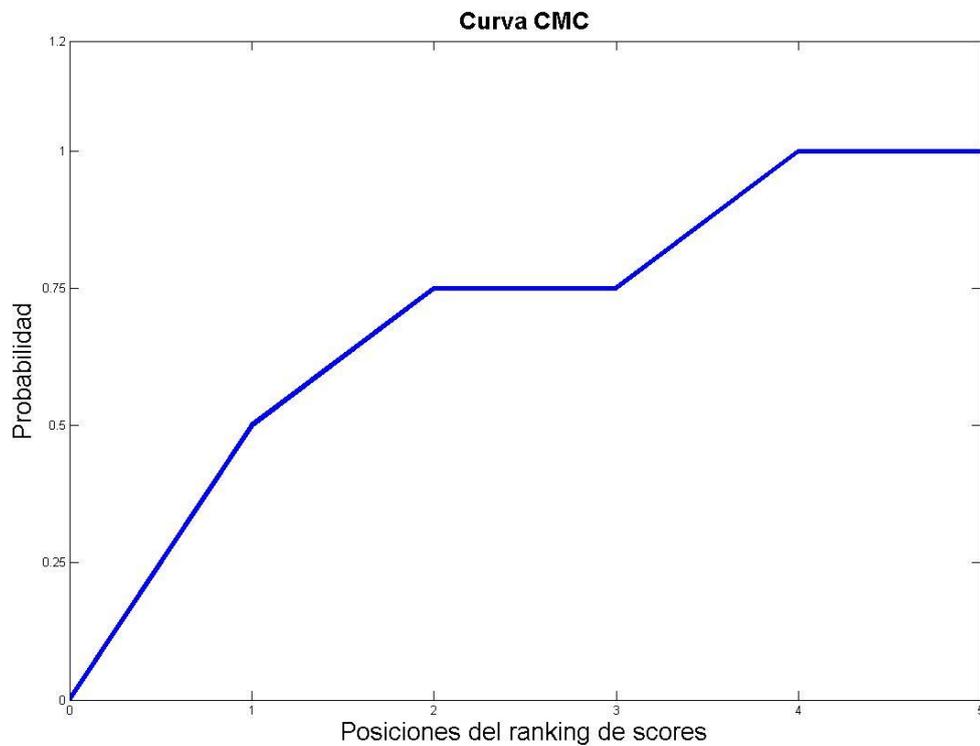


Figura 2-10: Ejemplo curva CMC

Una manera para mostrar dónde se está equivocando el sistema de reconocimiento son las matrices de confusión. Son matrices que tienen en la primera columna las muestras introducidos y en la primera fila los posibles modelos almacenados en el sistema con los que se les puede identificar. De tal manera que la casilla que representa el cruce de cada uno de estos caracteres se muestra la cantidad de veces que el sistema lo ha identificado con dicho modelo. En la figura 2-11 se observa un ejemplo de la matriz de confusión de la base de datos Firemaker, donde se puede observar que en la diagonal se muestra el mayor porcentaje de identificaciones, ya que las casillas de la diagonal son el cruce de las muestras introducidas con sus modelos acertados.

|   | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | A    | B    | C    | D    | E    | F    | G    | H    | I    | J    | K    | L    | M    | N    | O    | P    | Q    | R    | S    | T    | U    | V    | W    | Y    | Z    |   |
|---|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|---|
| 0 | 13   | 0    | 0    | 0    | 0    | 0    | 0.22 | 0    | 0    | 0    | 0    | 0    | 0.65 | 6.1  | 0    | 0    | 1.1  | 0    | 0.22 | 0    | 0    | 0    | 0    | 0    | 75   | 0    | 0    | 0    | 0    | 2.6  | 1.1  | 0    | 0    | 0.22 |      |   |
| 1 | 0    | 21   | 0    | 0    | 0    | 0    | 0.23 | 0    | 0    | 2.7  | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 75   | 0    | 0    | 0.23 | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0.46 | 0    | 0    | 0.23 | 0    |   |
| 2 | 0    | 0    | 41   | 0    | 0    | 0    | 0.45 | 0.45 | 0    | 0.89 | 0    | 0.45 | 0.45 | 0    | 0    | 0    | 0    | 0    | 0    | 0.89 | 0    | 1.8  | 0    | 0.45 | 0    | 0    | 1.8  | 0.89 | 0    | 0    | 0    | 0    | 0    | 0    | 0.51 |   |
| 3 | 0    | 0    | 0.42 | 67   | 0    | 2.1  | 0    | 0.84 | 0    | 0.42 | 0    | 2.1  | 0    | 0.84 | 0    | 0    | 0    | 0    | 0    | 0.42 | 1.3  | 0.42 | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0.42 | 2.1  |   |
| 4 | 0    | 0.5  | 0    | 0    | 56   | 0    | 0    | 0    | 0    | 2    | 1    | 0    | 2    | 0    | 0    | 0.5  | 0.5  | 2.5  | 0.5  | 0    | 0    | 2.5  | 0    | 0.5  | 1.5  | 0    | 3    | 0    | 1.5  | 0    | 15   | 2    | 0.5  | 8    | 0    |   |
| 5 | 0.64 | 0    | 0    | 0.64 | 0    | 27   | 0    | 0    | 0    | 0    | 0    | 0    | 0.64 | 0    | 0.64 | 0.64 | 1.9  | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0.64 | 0    | 2.6  | 0    | 63   | 1.3  | 0    | 0    | 0    | 0    | 0    |   |
| 6 | 0.47 | 0    | 0    | 0.47 | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 5.2  | 0.47 | 1.4  | 0    | 23   | 0    | 0.47 | 0    | 12   | 0    | 0    | 2.8  | 0    | 3.8  | 0    | 1.4  | 0    | 0.47 | 0.95 | 0    | 0    | 0    |      |   |
| 7 | 0    | 0.86 | 0    | 0    | 0    | 0    | 0    | 0    | 72   | 0    | 1.3  | 0    | 0    | 0    | 0    | 0    | 0.43 | 0    | 4.7  | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 1.3  | 3.4  | 0    | 9.1  | 0    | 1.3  | 0    | 3.9  | 2.2  |   |
| 8 | 0.58 | 0    | 1.2  | 0    | 0    | 2.3  | 0    | 0    | 22   | 0.58 | 0.58 | 2.9  | 0    | 1.7  | 2.3  | 0    | 0.58 | 0.58 | 1.7  | 1.2  | 0    | 0    | 0.58 | 0.58 | 12   | 0    | 4.7  | 0    | 41   | 0    | 0    | 0    | 0    | 3.5  |      |   |
| 9 | 0    | 0.54 | 0    | 1.1  | 6.5  | 0    | 0    | 2.7  | 0.54 | 42   | 0    | 0    | 0    | 1.1  | 0    | 2.2  | 0.54 | 1.1  | 0.54 | 0    | 0    | 0.54 | 0    | 1.6  | 0    | 11   | 0    | 11   | 0    | 0.54 | 0.54 | 0    | 17   | 0    |      |   |
| A | 0    | 0.59 | 0    | 0    | 0.15 | 0    | 0    | 0.03 | 0    | 0.03 | 95   | 0    | 0.03 | 0.21 | 0    | 0.03 | 0    | 0.53 | 0.06 | 0    | 0    | 0.12 | 0.15 | 0.97 | 0.12 | 0.12 | 1.3  | 0.09 | 0.03 | 0    | 0    | 0    | 0    | 0.12 | 0.03 |   |
| B | 0.39 | 0.26 | 0.13 | 1.2  | 0    | 0.26 | 1.8  | 0    | 1.4  | 0.26 | 1.3  | 54   | 0.92 | 6.7  | 3    | 0    | 5.6  | 0.66 | 0.39 | 0.13 | 0    | 0.92 | 0.66 | 0.26 | 3.9  | 0    | 4.1  | 1    | 6    | 0.13 | 0.26 | 0.13 | 0    | 0.13 | 3.8  |   |
| C | 0    | 0    | 0    | 0    | 0    | 0    | 0.55 | 0    | 0    | 0    | 0    | 90   | 0    | 2.6  | 0    | 0.37 | 0    | 0.18 | 0    | 6    | 0    | 0    | 0.37 | 0    | 0.37 | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    |   |
| D | 3    | 0    | 0    | 0    | 0    | 0    | 0.24 | 0    | 0    | 0    | 0.33 | 0.14 | 0.19 | 70   | 0    | 0.05 | 0.09 | 0.05 | 0.14 | 0.05 | 0    | 1    | 0    | 0.99 | 18   | 1.8  | 2    | 0.05 | 0.38 | 0    | 0.09 | 0.66 | 0    | 0    | 0    |   |
| E | 0    | 0    | 0    | 0    | 0    | 0    | 0.23 | 0    | 0    | 0    | 0.03 | 5.4  | 0    | 90   | 0.81 | 0.07 | 0    | 0.23 | 0    | 0.03 | 1.3  | 0    | 0    | 0.13 | 0    | 0.88 | 0.07 | 0.1  | 0.07 | 0    | 0    | 0    | 0    | 0    | 0.2  |   |
| F | 0    | 0.78 | 0    | 0    | 0.26 | 0    | 0    | 0.78 | 0    | 0    | 0.26 | 0    | 0    | 1.3  | 68   | 0    | 0    | 1.3  | 0    | 0.52 | 0.78 | 0    | 0    | 0    | 1.6  | 2.3  | 0    | 0    | 22   | 0    | 0    | 0    | 0    | 0    |      |   |
| G | 3    | 0    | 0    | 0    | 0.7  | 0    | 6.3  | 0    | 0    | 0.35 | 0.17 | 0    | 2.6  | 0    | 0.87 | 0    | 58   | 0    | 0    | 0.35 | 0    | 0.52 | 0    | 17   | 0    | 5.6  | 0    | 2.8  | 0    | 1.4  | 0.35 | 0    | 0.17 | 0    |      |   |
| H | 0    | 0    | 0    | 0    | 1.4  | 0    | 0    | 0.14 | 0    | 0.14 | 1.4  | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 75   | 0    | 0    | 0.54 | 0.27 | 13   | 3.1  | 0    | 0.14 | 3.3  | 0.14 | 0    | 1.4  | 0.82 | 0.14 | 0    | 0 |
| I | 0.05 | 16   | 0.1  | 0.1  | 0    | 0    | 0    | 0.1  | 0    | 0    | 0    | 0.05 | 0    | 0.1  | 0.1  | 0    | 0    | 0    | 81   | 0.52 | 0    | 0.57 | 0    | 0    | 0    | 0.05 | 0    | 0.1  | 0.63 | 0    | 0    | 0    | 0    | 0.05 |      |   |
| J | 0    | 0.79 | 0    | 0    | 0    | 0    | 0    | 0.79 | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 7.9  | 72   | 0    | 0    | 0    | 0    | 0    | 10   | 0    | 7.1  | 0    | 0.79 | 0    | 0    | 0.79 | 0    |   |
| K | 0    | 0.46 | 0    | 0    | 0.12 | 0    | 0.12 | 0    | 0    | 0.23 | 0    | 0.69 | 0    | 0.12 | 0.23 | 0    | 0.35 | 1.3  | 0.12 | 7.3  | 8.2  | 0.23 | 0.23 | 0    | 0.12 | 3.7  | 2    | 0    | 0    | 3.3  | 4.5  | 0.23 | 0.69 | 0.12 |      |   |
| L | 0    | 0.44 | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 2.2  | 0    | 0.27 | 0    | 0    | 0.88 | 0    | 0    | 0    | 0.88 | 0    | 0    | 0    | 96   | 0    | 0    | 0    | 0    | 0    | 0.35 | 0    | 0    | 0.18 | 0    | 0    |   |
| M | 0    | 0.14 | 0    | 0    | 0.34 | 0    | 0    | 0.14 | 0    | 0    | 0.27 | 0    | 0    | 0.27 | 0    | 0    | 8.2  | 0.07 | 0.14 | 0    | 0.07 | 78   | 6.2  | 0    | 0.07 | 4.3  | 0    | 0    | 0.41 | 0.69 | 0.48 | 0.27 | 0.27 | 0    | 0    |   |
| N | 0.02 | 0.04 | 0    | 0    | 0    | 0.07 | 0    | 0    | 0    | 0.46 | 0    | 0.09 | 0    | 0.02 | 0    | 0.75 | 0.27 | 0    | 0    | 0.32 | 0.55 | 95   | 0.11 | 0.09 | 1.2  | 0.02 | 0    | 0.02 | 0.07 | 0.75 | 0.39 | 0.05 | 0    | 0    | 0.05 |   |
| O | 12   | 0    | 0    | 0    | 0    | 0    | 0.18 | 0    | 0    | 0.04 | 0    | 0.29 | 3.7  | 0.07 | 0    | 1.1  | 0    | 0.04 | 0    | 0    | 0    | 0    | 0    | 0    | 0.04 | 80   | 0    | 0.81 | 0    | 0.11 | 0    | 1    | 0.59 | 0    | 0    |   |
| P | 0.26 | 0    | 0    | 0    | 0    | 0    | 0.26 | 0    | 0    | 0    | 0    | 0    | 6.9  | 0.26 | 3.1  | 0    | 0    | 0    | 0    | 0    | 0    | 0.51 | 0.26 | 1.5  | 1.5  | 74   | 8.5  | 0.77 | 0    | 1.3  | 0    | 0.51 | 0    | 0    |      |   |
| Q | 4.6  | 0.46 | 0    | 0    | 0.93 | 0    | 0.46 | 0    | 0    | 0.93 | 0    | 0.46 | 0.93 | 0    | 4.2  | 0    | 0.93 | 0.93 | 0    | 0.46 | 0    | 1.4  | 30   | 0.93 | 41   | 0.93 | 0.46 | 0.46 | 7.4  | 1.9  | 0.46 | 0    | 0    | 0    |      |   |
| R | 0.04 | 0.14 | 0.14 | 0    | 0.04 | 0    | 0    | 0.07 | 0    | 0    | 0.8  | 0.04 | 1.8  | 0.45 | 1.1  | 0.14 | 0.07 | 0.24 | 0.21 | 0.04 | 2.5  | 0.52 | 0.07 | 0.9  | 0.17 | 1.7  | 1.3  | 86   | 0    | 0.04 | 0.07 | 0.45 | 0.04 | 0.07 | 0.38 |   |
| S | 0    | 0    | 0    | 0.12 | 0.12 | 1.6  | 0.25 | 0    | 0.06 | 0    | 0.06 | 0.06 | 0.19 | 0.06 | 0    | 0.25 | 0    | 0.5  | 1.5  | 0    | 0    | 0    | 0    | 0    | 0.06 | 2.2  | 0    | 93   | 0.12 | 0    | 0.06 | 0    | 0    | 0    | 0    |   |
| T | 0    | 0.38 | 0    | 0    | 0    | 0    | 0    | 0.7  | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0.32 | 0    | 0    | 1    | 0    | 0    | 0.06 | 0    | 0.06 | 0    | 0.7  | 0.06 | 0    | 0.97 | 0    | 0    | 0    | 0    | 0.06 | 0    |   |
| U | 0.2  | 0    | 0    | 0    | 0.59 | 0    | 0.08 | 0    | 0    | 0    | 0.04 | 0    | 0    | 0    | 0    | 0.04 | 0.31 | 0.04 | 0.16 | 0.28 | 0.04 | 0.04 | 0.16 | 1.1  | 0    | 1.1  | 0    | 0    | 0    | 0    | 0.90 | 5.8  | 0.2  | 0.04 | 0    |   |
| V | 0.05 | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0.05 | 0    | 0    | 0    | 0    | 0.16 | 0.21 | 0.05 | 0.21 | 0.05 | 0.16 | 0.32 | 0    | 0.84 | 0    | 0    | 5.6  | 91   | 0.26 | 0.79 | 0    | 0    |      |   |
| W | 0    | 0.2  | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0.1  | 0    | 0    | 0    | 0    | 0    | 0    | 0.1  | 0.2  | 0    | 0.1  | 0    | 0    | 7.3  | 0.3  | 0    | 6.1  | 0    | 0    | 1.3  | 2.7  | 82   | 0    | 0    | 0    |      |   |
| Y | 0    | 0    | 0    | 0    | 4.6  | 0    | 0    | 0.44 | 0    | 3.3  | 0.44 | 0.22 | 0    | 0    | 0    | 0.22 | 0.44 | 1.5  | 0.44 | 0.22 | 0    | 0.44 | 0    | 0.44 | 0.22 | 7.9  | 0    | 0.44 | 0.66 | 0.44 | 4.4  | 0.22 | 73   | 0    | 0    |   |
| Z | 0    | 0.43 | 13   | 0    | 0    | 0    | 0    | 1.6  | 0    | 0    | 0.22 | 0    | 0.76 | 0    | 0.65 | 0.11 | 0    | 0    | 2.4  | 0.33 | 0    | 2    | 0    | 0    | 0    | 0.11 | 3.3  | 0.43 | 0    | 0.33 | 0    | 0    | 0.11 | 75   | 0    |   |

Figura 2-11: Ejemplo de matriz de confusión (Firemaker)



# 3 Adquisición y pre-procesado de la base de datos

---

En esta sección se explicará el proceso de adquisición y pre-procesado de la base de datos de este proyecto.

La adquisición se realizó con un software llamado Biógrafo [27] que permitía recortar manualmente y almacenar los caracteres manuscritos de todos los documentos escaneados. La base de datos a adquirir está diferenciada en subcorpus:

- Indubitados
- Dubitados

Una vez realizada la adquisición, como tarea previa a la obtención de resultados experimentales de reconocimiento de caracteres, se pre-procesó la base de datos obtenida en Biógrafo para obtener las características necesarias que la herramienta de reconocimiento de carácter requería.

## 3.1 Adquisición

La adquisición de la base de datos fue una tarea manual realizada durante 2 años en colaboración con el Ministerio del Interior de España, quienes prestaron sus documentos manuscritos de 478 personas oficialmente identificadas (indubitados) y 49 personas identificadas por un biógrafo profesional (dubitados).

La tarea de este proyecto comenzó a partir de los documentos ya escaneados y almacenados en un disco duro, estaban divididos en dos grupos:

- Individuos Indubitados: textos procedentes de escritores identificados. Cuyos caracteres formarán el conjunto de entrenamiento.
- Individuos Dubitados: textos procedentes de escritores identificados visualmente por un biógrafo. Cuyos caracteres formarán el conjunto de test.

De cada uno de los 478 indubitados hay tres tipos de documentos manuscritos:

- En minúsculas
- En mayúsculas
- Números

Los dos primeros consisten en un texto manuscrito con gran variedad de caracteres alfabéticos. El tercer documento consiste en una hoja con operaciones matemáticas realizadas a mano por cada indubitado.

En las Figuras 3-1 a 3-3 se pueden observar un ejemplo de un fragmento de cada uno de estos tres tipos de documentos.

corriendo para que no los cogieran.  
 En la carrera llegaron a topetar  
 incluso con un grupo de [REDACTED]  
 que andaba de vista nocturna  
 por la ciudad y que desconocía  
 la existencia de la [REDACTED]  
 por estos lugares.

Figura 3-1: Fragmento de un documento manuscrito por un individuo indubitado en minúsculas

DESTACAN LAS CUMBRES DE [REDACTED]  
 DONDE ENCONTRAMOS EJEMPLARES  
 DE LA [REDACTED] Y EL ZORRO.  
 LA RIA DE [REDACTED] CONSTITUTE  
 LA ZONA HUMEDA MAS IMPORTANTE  
 DE [REDACTED], DE 12 KM.s DE LONGITUD,

Figura 3-2: Fragmento de un documento manuscrito por un individuo indubitado en letras mayúsculas

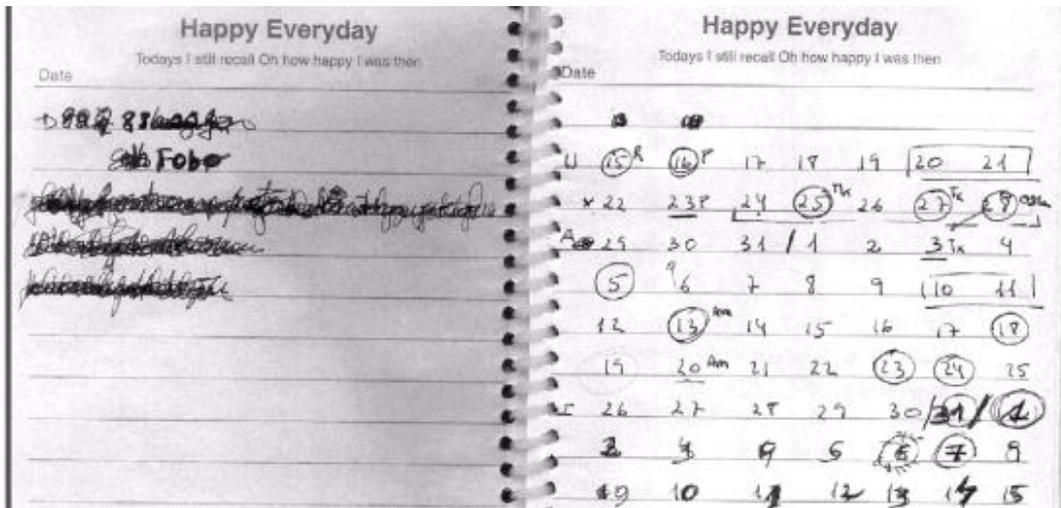
|   |   |
|---|---|
| $  \begin{array}{r}  30734946 \\  + 12793040 \\  \hline  43527986  \end{array}  $ | $  \begin{array}{r}  6215763 \\  + 9842654 \\  \hline  16058417  \end{array}  $   |
| $  \begin{array}{r}  50780249 \\  - 41365574 \\  \hline  09415675  \end{array}  $ | $  \begin{array}{r}  90123768 \\  - 75318596 \\  \hline  14805172  \end{array}  $ |

Figura 3-3: Fragmento de un documento manuscrito por un individuo indubitado realizando operaciones matemáticas

Los documentos de los dubitados fueron incautados, por lo que su texto y formato es muy variable. En estos documentos la cantidad de caracteres obtenida es mucho menor, debido a que en estos no existe tanta variedad de caracteres y a que la cantidad de dubitados es mucho menor que la de indubitados.

En la Figura 3-4 se muestra un fragmento de texto de un dubitado. Como podemos observar, el tipo de documento varia mucho de un dubitado a un indubitado; en el caso de

los primeros se trata de documentos escritos sobre cualquier tipo de papel y con cualquier utensilio de escritura, por lo que el conjunto de test creado a partir de estos documentos tiene mucha mayor variabilidad.



**Figura 3-4: Ejemplo de documento manuscrito de un individuo dubitado.**

Para la adquisición de la base de datos se buscó recoger 5 muestras de cada caracter escritas por cada individuo. Aún así no se pudo obtener esta cantidad de muestras en todos los casos debido a que ciertos caracteres eran más escasos en algunos de los documentos.

En la sección 4 se muestra la caracterización de la base de datos tras su adquisición, donde se puede observar la cantidad de muestras de cada caracter. Para caracteres como “X”, “W”, “K”, “x”, “w” ó “k”, la cantidad de muestras no es tan abundante como en el resto de los casos, al no ser letras tan comunes en el alfabeto español.

### 3.1.1 Herramienta Biógrafo

La adquisición de la base de datos se realizó con un software llamado Biógrafo que integra una herramienta para la adquisición de caracteres [27].

Biógrafo es una aplicación de cliente programada en JavaScript que presenta dos modos de operar:

1. Modo local: tanto la aplicación cliente como la base de datos funcionan en la misma máquina.
2. Modo remoto: alguna de las copias de la aplicación cliente instaladas en diferentes máquinas se comunican con una única base de datos instalada en un servidor remoto.

Para este proyecto se utilizó la herramienta Biógrafo en modo remoto.

Biógrafo tiene dos funcionalidades principales:

1. Adquisición manual de caracteres de documentos manuscritos de individuos indubitados y/o dubitados y la organización de los datos adquiridos.
2. Función de identificación automática de individuos: a partir de las muestras adquiridas en Biógrafo, la herramienta crea una lista de posibles individuos indubitados para cada individuo dubitado.

Para este proyecto se utilizó la herramienta de adquisición de datos de Biógrafo. La segmentación manual de la base de datos fue una tarea de dos años de duración. A continuación, en la figura 3-5, se muestra un pantallazo de la herramienta utilizada para la segmentación manual de los documentos manuscritos.

La herramienta de adquisición incorpora 3 menús principales llamados: Individuos, Documentos y Muestras. En el menú Individuos se registran en la base de datos los nuevos individuos indubitados o dubitados, asignándoseles un número de identificación. En el menú de Documentos, se importa la documentación digitalizada asignándose a un individuo previamente registrado. Por último, en el menú de Muestras se pueden adquirir las muestras de cada individuo.

Biógrafo permite adquirir muestras de los 62 distintos caracteres posibles; 26 letras mayúsculas, 26 letras minúsculas y 10 números.

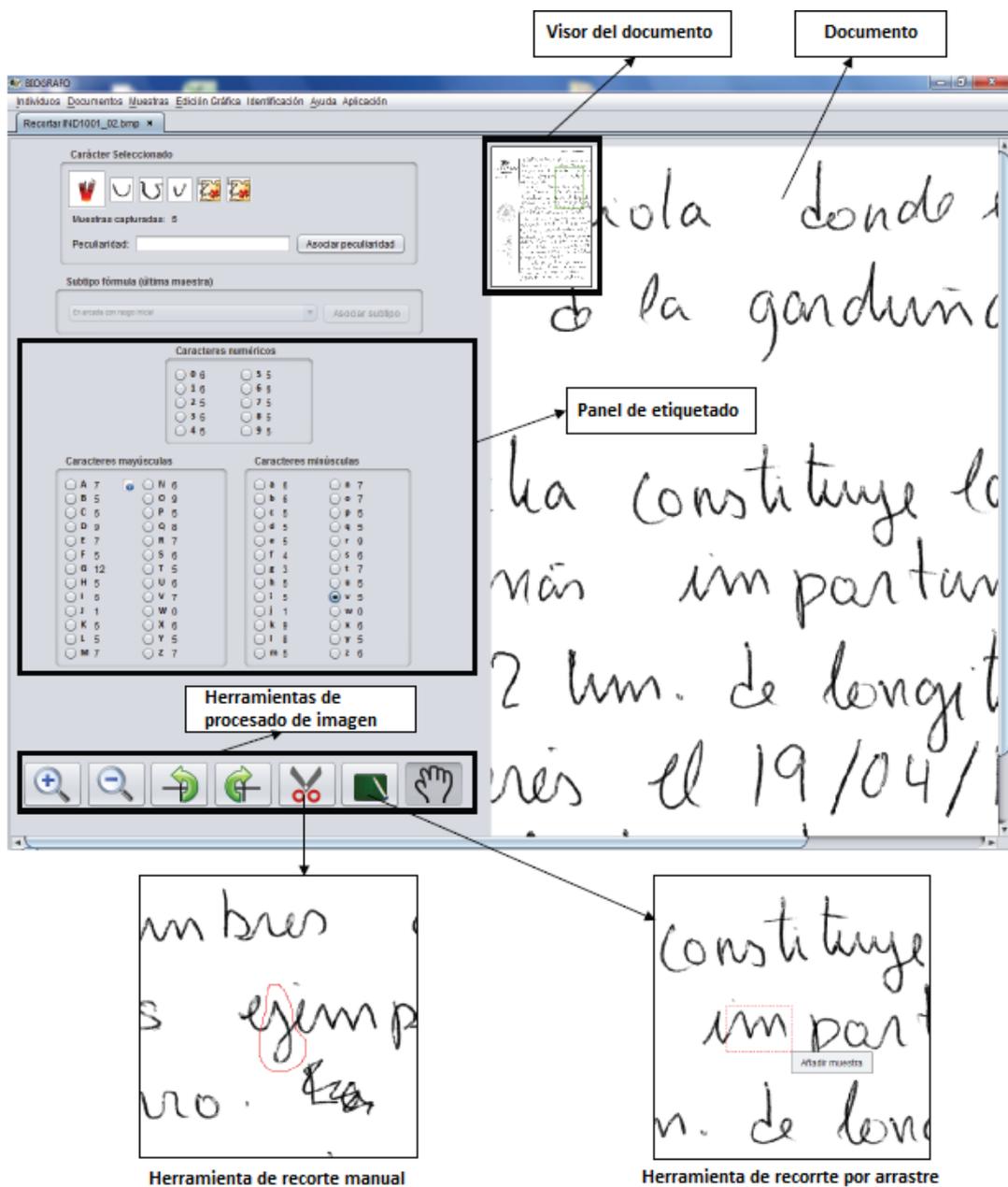


Figura 3-5: Pantallazo de la herramienta de adquisición de muestras de Biógrafo

En la adquisición de la base de datos del proyecto, se realizaron los siguientes pasos:

- Como primer paso, se registraron los 478 indubitados y los 49 dubitados en el menú de Identificación de la herramienta Biógrafo.
- Para los 478 indubitados se asignaron las etiquetas desde IND\_1001 hasta IND\_1488 (hay más etiquetas que indubitados porque 10 de ellos dieron problemas al adquirir la base de datos, pero se dejó sus etiquetas por cuestiones de organización). De una manera similar se hizo para los 49 dubitados, asignándoles las etiquetas DUB\_1001 hasta DUB\_1049.
- A continuación, se introdujeron los respectivos documentos de cada individuo a través del menú de Documentación, enlazando cada documento a su respectivo individuo.
- Por último se procedió a la adquisición de las muestras de cada documento a través de la herramienta Muestras.

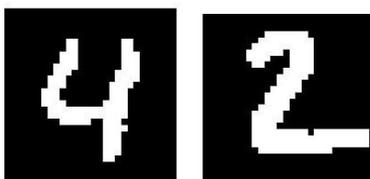
### **3.2 Pre-procesado**

Una vez adquirida la base de datos se pasó al procesado de las imágenes para normalizarlas y así poder proceder a la obtención de resultados de referencia.

La primera decisión que se tomó fue procesar las imágenes de la base de datos adquirida con Biógrafo de tal manera que tuviesen el mismo formato que la base de datos de caracteres manuscritos referente en el ámbito de investigación, MNIST [1]. De esta manera nuestra base de datos será más fácil de utilizar para futuros experimentos con diferentes herramientas de detección de caracteres manuscritos, ya que muchas de estas herramientas están adaptadas a las características de la base de datos MNIST.

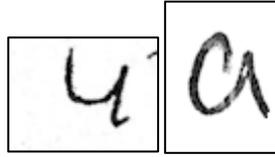
Las características de la base de datos de MNIST son las siguientes:

- Imágenes en escala de grises, 256 niveles de gris, desde 0 (negro) a 255 (blanco).
- Imágenes normalizadas en una imagen de 20x20 píxeles preservando su relación de aspecto.
- El centro de masas de las imágenes 20x20 están en el centro geométrico de la imagen 28x28.
- Estas imágenes están almacenadas en matrices de vectores.



**Figura 3-6: Ejemplo de imágenes procesadas de la base de datos MNIST**

Se extrajeron todas las imágenes de caracteres manuscritos adquiridas con Biógrafo. El aspecto de imágenes sin procesar de la base de datos de Biógrafo se puede apreciar en la Figura 3-7.



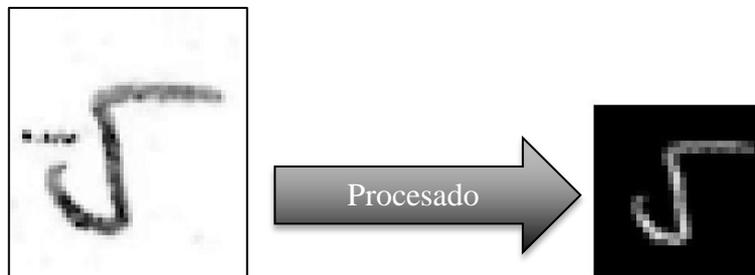
**Figura 3-7: Ejemplo de imágenes sin procesar de la base de datos Biógrafo**

Las imágenes de caracteres manuscritos de Biógrafo están en formato imagen de mapa de bits (.bmp). Cada imagen tiene dimensiones distintas y muchas de ellas necesitan ser procesadas para obtener una imagen más precisa del carácter, ya que contienen trazos de caracteres adyacentes, manchas o errores de escritura.

El procesamiento de la base de datos Biógrafo se realizó con Matlab, una herramienta de software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). En las siguientes secciones se explicarán los pasos seguidos en el pre-procesado de este proyecto:

- Binarización de las imágenes, con el objetivo de poder utilizar algunas funciones de Matlab que trabajan con imágenes binarias.
- Segmentación de caracteres, para eliminar de la imagen todo lo que no perteneciese al carácter en cuestión.
- Normalización de las imágenes, para obtener las dimensiones adecuadas en todas las imágenes de la base de datos.

En la figura 3-8 se muestra la evolución del procesamiento de imágenes para la base de datos de Biógrafo.



**Figura 3-8: Procesado de inicio a fin**

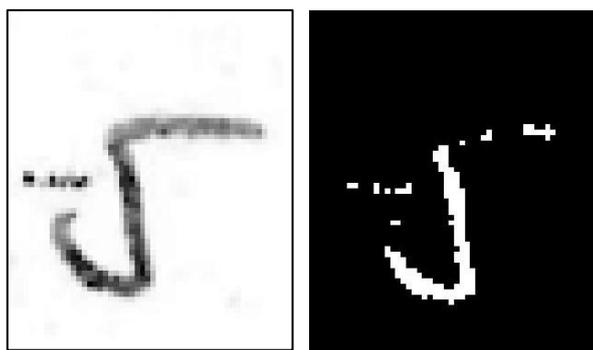
Las imágenes pre-procesadas se almacenan en matrices de vectores (elaboradas con Matlab), al igual que MNIST, ya que la herramienta de detección que se va a utilizar, requiere las imágenes de caracteres almacenadas en forma de vector.

### **3.2.1 Binarización de las imágenes**

Como primer paso se procedió a la tarea de umbralización con el fin de binarizar las imágenes. Esta necesidad nace a partir de que las herramientas a usar para el procesamiento de las mismas trabajan con imágenes binarias.

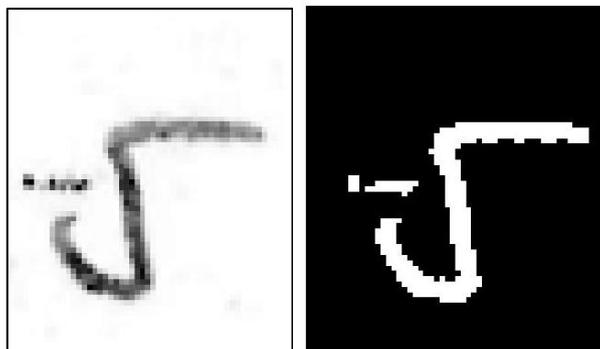
El valor de umbralización se decidió tras observar el histograma de las imágenes de Biógrafo y realizar varias pruebas con diferentes valores para el umbral. En dicho histograma se podían apreciar perfectamente dos picos: uno alrededor de 255 (el fondo de la imagen) y otro alrededor de 0 (el carácter).

En primer lugar se decidió establecer el valor de umbral en el punto medio entre los dos picos del histograma, 128. Estableciendo este valor de umbral se observó que no se tenían en cuenta algunos píxeles del caracter debido a que el escritor ejercía menor presión con el bolígrafo o al mal funcionamiento del bolígrafo. De esta forma aparecen píxeles más claros que también pertenecen al caracter pero que no se tienen en cuenta con este valor de umbral. En el ejemplo de la figura 3-9 se puede observar este efecto, donde a la izquierda se presenta un 5 de la base de datos Biógrafo y a la derecha su correspondiente binarización. En este ejemplo se muestran algunos trazos del caracter de la base de datos Biógrafo en un tono más suave. Al binarizar dicha imagen, no se tienen en cuenta los píxeles más suaves del caracter debido al valor seleccionado para el umbral.



**Figura 3-9: Imagen de la base de datos Biógrafo (carácter manuscrito 5) y su correspondiente binarización con valor de umbral 128**

Por consiguiente, tras varios experimentos y observaciones, se decidió escoger un valor de umbral más elevado, 200, de tal forma que la imagen binarizada contemplase estos píxeles más suaves (ver figura 3-10).



**Figura 3-10: Imagen de la base de datos Biógrafo (carácter manuscrito 5) y su correspondiente binarización con valor de umbral 200**

### 3.2.2 Segmentación de caracteres

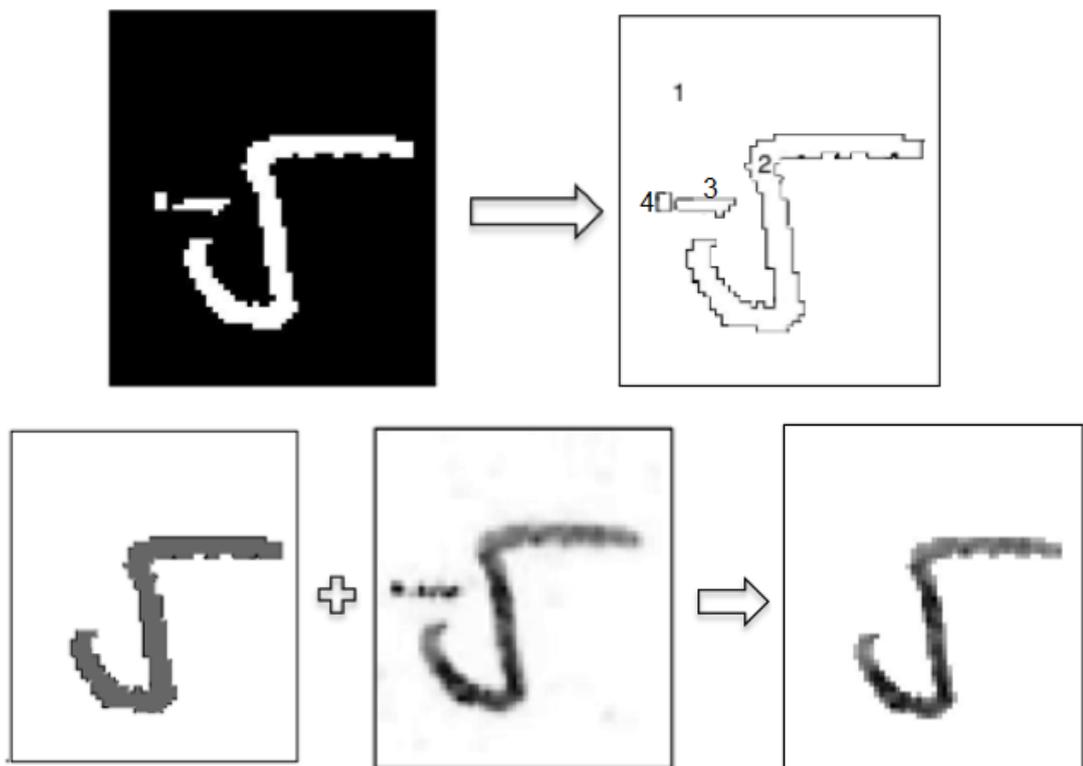
Tras la binarización de las imágenes de la base de datos de Biógrafo, se procedió a la eliminación de trazos espurios que no pertenecían al caracter (como pueden ser líneas de caracteres cercanas fruto de errores humanos al segmentar los caracteres).

El método utilizado para la eliminación de estos trazos en la imagen no pertenecientes al caracter consistió en localizar los objetos conectados en la imagen [28] y obtener el

segundo objeto más grande [29], ya que el primer objeto es el fondo de imagen y el segundo objeto más grande el carácter en sí, siendo el resto de objetos trazos que deben ser eliminados. Para obtener un resultado adecuado en este proceso hacía falta que la imagen fuese binaria.

Dado que el objetivo final es tener imágenes en escala de grises como en MNIST, se aprovecha la eliminación de trazos espurios para además convertir la imagen binaria en escala de grises. Esto se realiza utilizando los trazos obtenidos en la localización de objetos como una máscara, de tal manera que indexando la imagen original (en escala de grises) y la máscara de objetos (con los objetos deseados), se obtiene el carácter en sus valores originales y el resto de la imagen en blanco, como fondo de imagen.

La figura 3-12 ilustra lo explicado previamente. Primero la imagen binarizada, posteriormente los 4 objetos detectados de esta imagen. A continuación cogemos el primer objeto (marcado como 1, fondo de la imagen) y el segundo objeto (marcado como 2, carácter deseado), eliminando el resto de objetos (marcados como 3 y 4). Se aplica la máscara creada a partir de la eliminación de objetos no deseados a la imagen original, para así obtener la imagen deseada, sin trazos espurios y en escala de grises.



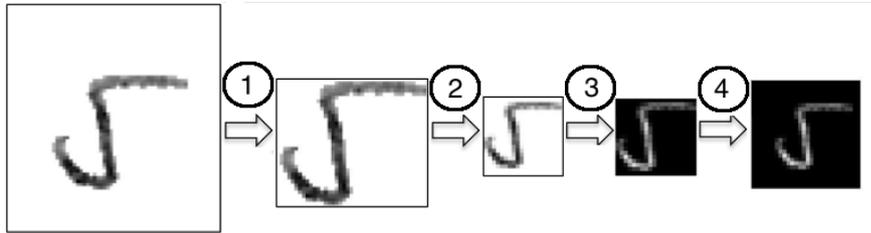
**Figura 3-11: Proceso de segmentación y conversión a escala de grises.**

### 3.2.3 Normalización de las imágenes

Tras la segmentación de las imágenes, sin trazos espurios, se procede a la normalización para obtener las mismas medidas que en MNIST.

La normalización de las imágenes se realizó mediante los pasos que se muestran en la figura 3-12:

1. Detección del caracter en la imagen, para posteriormente poder calcular su relación de aspecto al cambiar de tamaño.
2. Normalización de todas las imágenes a un tamaño de 20x20 píxeles manteniendo su relación de aspecto.
3. Se invirtieron los valores de los píxeles para tener el fondo oscuro y el caracter en color claro, como las imágenes de MNIST.
4. Calculo del centro de masas del caracter [30] para colocarlo en el centro geométrico de un cuadrado 28x28 píxeles.



**Figura 3-12: Normalización paso a paso**



## **4 Descripción de la base de datos**

---

A continuación se mostrarán en detalle las principales características de la base de datos adquirida con la herramienta Biógrafo así como las características de base de datos tras el pre-procesado para la obtención de resultados de referencia.

La base de datos tiene distintas características antes y después del procesado debido a que durante esta tarea se descartaron las muestras que se consideraron defectuosas por fallos en la adquisición o por contener demasiados trazos espurios en la imagen.

### **4.1 Base de datos adquirida con biógrafo**

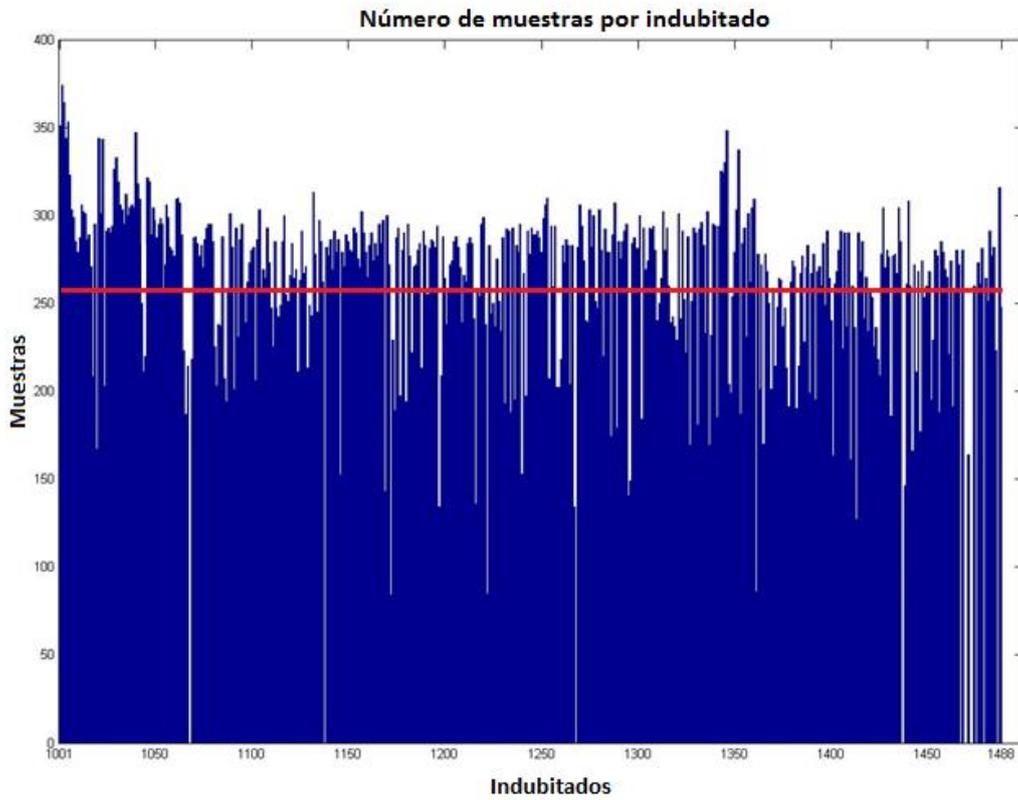
La tarea de adquisición de caracteres dio lugar a una base de datos muy amplia y con una gran variedad. Contiene números, letras minúsculas y mayúsculas, procedente de documentos adquiridos bajo supervisión y documentos no supervisados.

La base de datos adquirida se dividió en dos grupos:

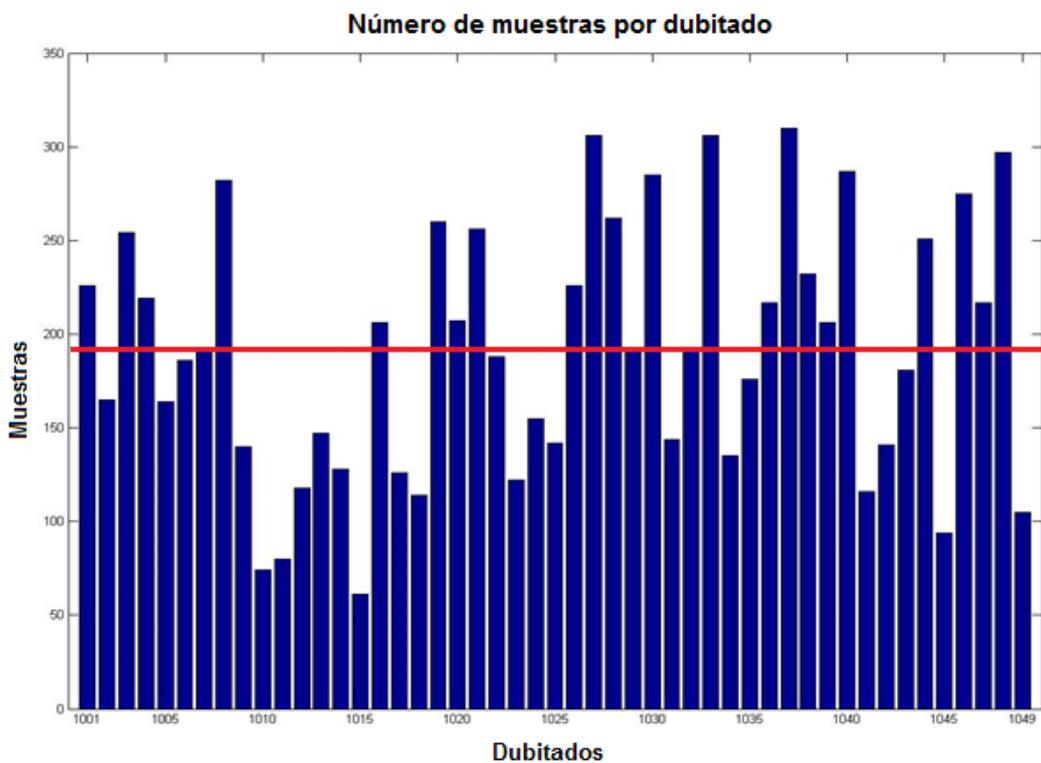
- Subcorpus indubitados.
- Subcorpus dubitados.

A continuación se van a mostrar las principales características de cada subcorpus de la base de datos, mediante unas gráficas ilustrativas:

- En la Figura 4-1 se observa el número de muestras que tiene cada indubitado. Se puede observar que cada individuo tiene una gran cantidad de muestras, siendo éstas el conjunto de números, letras minúsculas y mayúsculas. El número medio de muestras por cada indubitado es de 257. En esta figura se puede observar que hay algunos individuos que no tienen muestras, por lo que no han sido tenidos en cuenta en el proyecto.
- En la Figura 4-2 se indica el número de muestras por cada dubitado. Estas muestras son también el conjunto de números, letras minúsculas y mayúsculas. En este caso la cantidad de muestras por individuo es menor ya que cada dubitado tiene menos documentos manuscritos asignados y debido a que estos documentos no son un texto predeterminado con gran variedad de caracteres alfanuméricos, como en el caso de los documentos de indubitados. El número medio de muestras por cada dubitado es de 191, en comparación con las 257 de los indubitados.
- En la Figura 4-3 se representa el número de muestras que hay de cada caracter en el subcorpus indubitados. La media de muestras por caracter es 2026. Para los casos de caracteres como “w”, “x”, “y”, “W”, “X”, y “Y” la cantidad de muestras es muy inferior a la media debido a que estas letras no predominan en el alfabeto hispano.
- En la Figura 4-4 se observa el número de muestras que hay de cada caracter en el subcorpus dubitados. La media de muestras por caracter es 151, un valor notablemente menor que en el caso de indubitados, por las mismas razones ya descritas: hay menos documentos y estos tienen menor variedad de caracteres. De una forma similar que en el subcorpus indubitados, para los casos de caracteres como “f”, “q”, “w”, “y”, “Q”, y “W” la cantidad de muestras es muy inferior a la media.



**Figura 4-1: Número de muestras por cada individuo indubitado. La recta roja indica el número medio: 257.**



**Figura 4-2: Número de muestras por cada individuo dubitado. La recta roja indica el número medio: 191.**

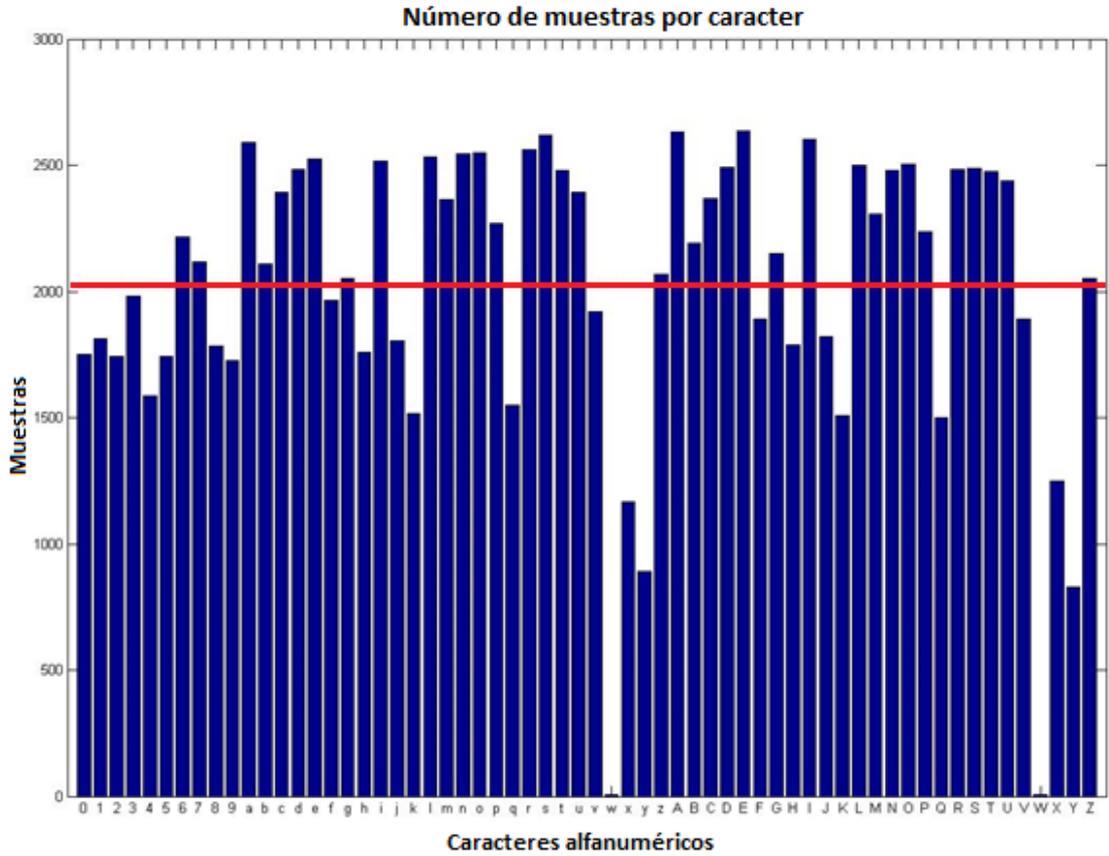


Figura 4-3: Número de muestras por caracter (subcorpus indubitados). La recta roja indica el número medio: 2026.

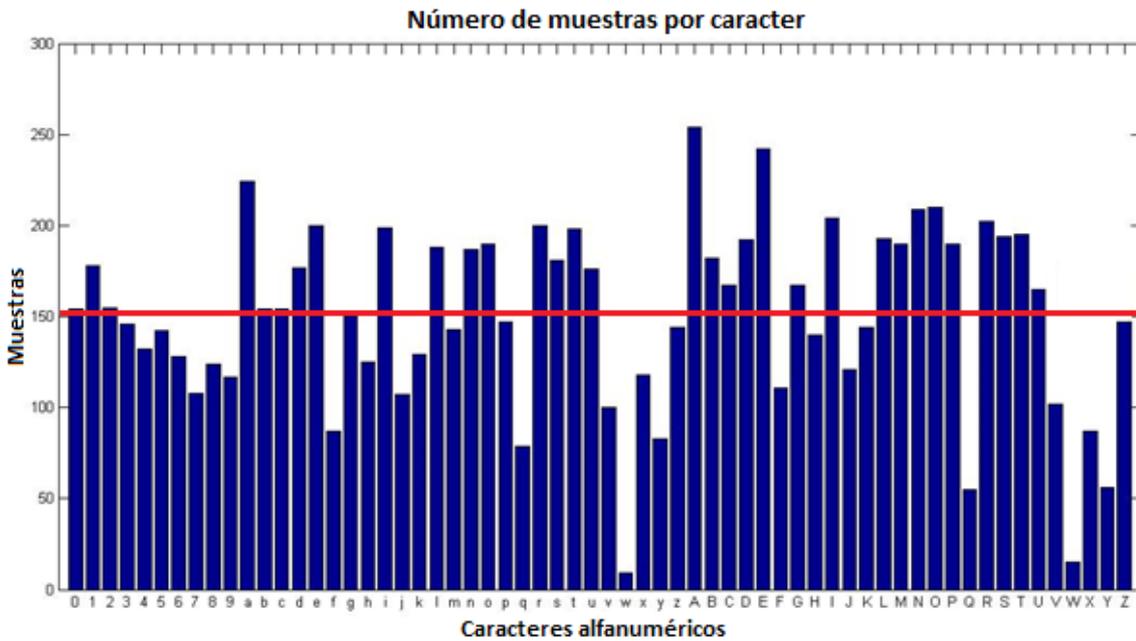


Figura 4-4: Número de muestras por caracter (subcorpus dubitados). La recta roja indica el número medio: 151.

## 4.2 Base de datos pre-procesada

Tras la adquisición de la base de datos con Biógrafo, se procedió al pre-procesado de las imágenes y la organización en matrices de Matlab. Durante esta tarea se desecharon varias muestras que se detectaron defectuosas por fallos humanos en la tarea de segmentación manual.

En las tablas 4-5 a 4-10 se muestra la cantidad de imágenes que tiene cada caracter según el tipo de escritor. Los caracteres 'w' y 'W' no se consideran para la obtención de resultados, debido a que la cantidad de muestras disponible era demasiado baja (Figuras 4-3 y 4-4) hecho que sesgaría los resultados. Además, en estas tablas de indubitados se marcan en amarillo los valores más bajos en términos de cantidad, ya que este hecho será importante en el protocolo planteado para la obtención de los resultados de referencia.

| IND | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    |
|-----|------|------|------|------|------|------|------|------|------|------|
|     | 1730 | 1780 | 1695 | 1932 | 1418 | 1635 | 2192 | 2077 | 1741 | 1661 |

**Tabla 4-5: Cantidad de muestras por caracter numérico del conjunto indubitado**

| IND | A    | B    | C    | D    | E    | F    | G    | H    | I    | J    | K    | L    |      |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|
|     | 2552 | 2074 | 2241 | 2662 | 1852 | 1437 | 1932 | 1451 | 2355 | 1630 | 1332 | 2347 |      |
|     | M    | N    | O    | P    | Q    | R    | S    | T    | U    | V    | X    | Y    | Z    |
|     | 2171 | 2301 | 2405 | 2040 | 1395 | 2284 | 2286 | 1946 | 2286 | 1802 | 1155 | 719  | 1913 |

**Tabla 4-6: Cantidad de muestras por caracter del alfabeto en mayúsculas del conjunto indubitado**

| IND | a    | b    | c    | d    | e    | f    | g    | h    | i    | j    | k    | l    |      |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|
|     | 1754 | 1677 | 2079 | 1660 | 1824 | 1549 | 1556 | 1269 | 2277 | 1708 | 1059 | 2416 |      |
|     | m    | n    | o    | p    | q    | r    | s    | t    | u    | v    | x    | y    | z    |
|     | 1595 | 1876 | 1678 | 1724 | 1076 | 2013 | 1968 | 2011 | 1872 | 1624 | 1054 | 684  | 1649 |

**Tabla 4-7: Cantidad de muestras por caracter del alfabeto en minúsculas del conjunto indubitado**

| DUB | 0   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | 145 | 163 | 142 | 130 | 118 | 115 | 124 | 105 | 112 | 104 |

**Tabla 4-8: Cantidad de muestras por caracter numérico del conjunto dubitado**

| DUB | A   | B   | C   | D   | E   | F   | G   | H   | I   | J   | K   | L   |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | 214 | 164 | 143 | 155 | 167 | 88  | 142 | 122 | 165 | 114 | 119 | 157 |     |
|     | M   | N   | O   | P   | Q   | R   | S   | T   | U   | V   | X   | Y   | Z   |
|     | 156 | 174 | 165 | 162 | 52  | 170 | 161 | 145 | 145 | 107 | 70  | 54  | 131 |

**Tabla 4-9: Cantidad de muestras por caracter del alfabeto en mayúsculas del conjunto dubitado**

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| DUB | a   | b   | c   | d   | e   | f   | g   | h   | i   | j   | K   | l   |     |
|     | 184 | 145 | 131 | 161 | 172 | 92  | 135 | 125 | 164 | 100 | 120 | 169 |     |
|     | m   | n   | o   | p   | q   | r   | s   | t   | u   | v   | X   | y   | z   |
|     | 132 | 169 | 171 | 140 | 79  | 170 | 170 | 181 | 152 | 94  | 101 | 76  | 133 |

**Tabla 4-10: Cantidad de muestras por caracter del alfabeto en minúsculas del conjunto dubitado**

Cabe destacar por último que en la base de datos pre-procesada se ha excluido la información sobre los escritores: sólo se almacenan los caracteres, y si estos proceden de documentos indubitados o dubitados.



## 5 Resultados de referencia

---

Una vez procesada toda la base de datos y almacenada en matrices de Matlab, se procedió a la obtención de resultados de referencia con un sistema de reconocimiento de caracteres. Para ello se propuso un protocolo experimental en el que se fijaron y organizaron los experimentos a realizar con la base de datos creada.

En esta sección se explicará el protocolo seleccionado, así como el sistema de reconocimiento utilizado, para finalmente mostrar los resultados experimentales obtenidos para los diferentes escenarios planteados.

### 5.1 Protocolo experimental

Como primer paso del protocolo experimental se distinguieron las diferentes pruebas que se querían realizar con la amplia base de datos procesada que se había conseguido (detallada en la sección 4.2).

Se diferencian dos tipos de imágenes de caracteres manuscritos según el tipo de escritor:

- Indubitados.
- Dubitados.

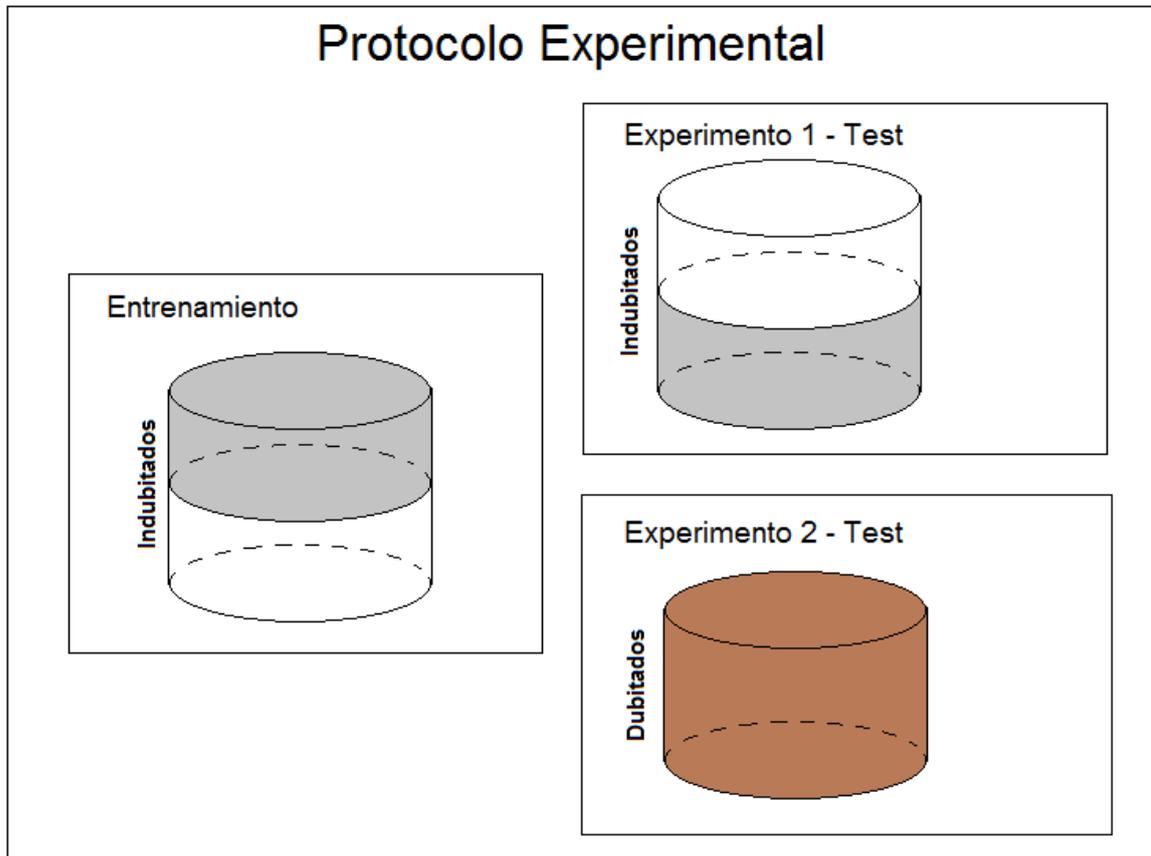
Ambos tipos de escritores tienen imágenes de caracteres manuscritos del tipo:

- Números.
- Mayúsculas.
- Minúsculas.

Según los datos mostrados en la sección 4.2 se proponen dos experimentos:

- Experimento 1: indubitados vs indubitados. Se usa una parte de las imágenes de indubitados para el conjunto de entrenamiento y la otra parte de imágenes de indubitados para el conjunto de test. De este experimento se obtienen 4 escenarios:
  - Números: solo imágenes de caracteres numéricos.
  - Mayúsculas: solo imágenes de caracteres en letras mayúsculas.
  - Minúsculas: sólo imágenes de caracteres en letras minúsculas.
  - Todo: mezcla de imágenes de caracteres numéricos, mayúsculas y minúsculas.
- Experimento 2: dubitados vs indubitados. Usar imágenes de indubitados como conjunto de entrenamiento (el mismo conjunto de entrenamiento que se usa en el Experimento 1) y las imágenes de los dubitados como conjunto de test. Al igual que en el experimento anterior, se obtienen 4 escenarios:
  - Números: solo imágenes de caracteres numéricos.
  - Mayúsculas: solo imágenes de caracteres en letras mayúsculas.
  - Minúsculas: sólo imágenes de caracteres en letras minúsculas.
  - Todo: mezcla de imágenes de caracteres numéricos, mayúsculas y minúsculas.

Por lo tanto se han analizado 8 escenarios diferentes, 4 en cada experimento. En la figura 5-1 se muestra gráficamente los conjuntos de entrenamiento y test que se han utilizado para ambos experimentos.



**Figura 5-1: Conjuntos de entrenamiento y test. Protocolo experimental.**

Para analizar los resultados de los experimentos se usarán las siguientes medidas:

- Tasas de error de todas las evaluaciones: mide el porcentaje de ocasiones que el reconocedor no acierta en la evaluación de una muestra respecto al total de muestras evaluadas.
- Curvas CMC del escenario Todo en ambos experimentos: el sistema de reconocimiento crea una lista jerárquica asociando modelos al carácter evaluado. Esta lista tiene el carácter más similar reconocido en la primera posición y el menos parecido en la última posición. No siempre se acierta asociando el carácter evaluado a su correspondiente en la primera posición de la lista, pero puede ser que en segunda posición sí que esté el carácter acertado. La curva CMC (*Cumulative Match Curve*) mide la probabilidad de acierto acumulada en cada posición jerárquica de esta lista.
- Matriz de confusión del escenario Todo en ambos experimentos: la matriz de confusión muestra la relación de semejanza que tiene cada carácter consigo mismo y con el resto de caracteres de acuerdo al sistema de reconocimiento.

### 5.1.1 Protocolo Experimento 1

Para el Experimento 1 se utilizaron únicamente las imágenes de los indubitados. Se dividió el conjunto de imágenes en dos subconjuntos: entrenamiento y test.

El criterio que se decidió para crear estos dos subconjuntos fue que la cantidad de muestras por caracter en el conjunto de entrenamiento debía ser la misma para todos los caracteres. De esta manera los resultados futuros serían coherentes y rigurosos, debido a que la herramienta de detección de caracteres, entrenará así, a todos los caracteres en las mismas condiciones. Cabe destacar que la base de datos MNIST no tiene en cuenta esta particularidad de igualar la cantidad en el conjunto de entrenamiento.

Para elegir el número común de cada conjunto de entrenamiento se miran aquellos caracteres de cada tipo (números, mayúsculas y minúsculas) que tienen menos muestras. En las tablas 4-1 a 4-3 de indubitados están destacados en amarillo, estos son:

- Números: el caracter '4' con 1418 muestras.
- Mayúsculas: el caracter 'Y' con 719 muestras.
- Minúsculas: el caracter 'y' con 684 muestras.

Posteriormente se dividen entre dos estas cantidades mínimas. Por lo tanto, la cantidad que se tendrá, en cada evaluación del Experimento 1 es:

- Números: **709** muestras por caracter.
- Mayúsculas: **359** muestras por caracter.
- Minúsculas: **342** muestras por caracter.

La cantidad de muestras por caracter en el conjunto de test es variable y estará compuesta por todas las muestras que tiene cada caracter del conjunto de indubitado, quitándole las muestras que se usan para el conjunto de entrenamiento. En la Tabla 5-1 se puede observar la cantidad de muestras para cada evaluación del Experimento 1 en el conjunto de test.

En consecuencia, el escenario Todo (que tiene en cuenta números, letras mayúsculas y minúsculas) tendrá un tamaño máximo para el conjunto de muestras de entrenamiento igual al mínimo entre números, letras mayúsculas y minúsculas:

- Todo: **342** muestras por caracter.

Para dicho escenario se fusionan los conjuntos de test de números, letras mayúsculas y minúsculas, que se muestran en la Tabla 5-1.

| NÚMEROS  |          | MAYÚSCULAS |          | MINÚSCULAS |          |
|----------|----------|------------|----------|------------|----------|
| Caracter | Muestras | Caracter   | Muestras | Caracter   | Muestras |
| 0        | 1021     | A          | 2193     | a          | 1412     |
| 1        | 1071     | B          | 1715     | b          | 1335     |
| 2        | 986      | C          | 1882     | c          | 1737     |
| 3        | 1223     | D          | 1977     | d          | 1318     |
| 4        | 709      | E          | 1493     | e          | 1482     |
| 5        | 926      | F          | 1078     | f          | 1207     |
| 6        | 1483     | G          | 1573     | g          | 1214     |
| 7        | 1368     | H          | 1092     | h          | 927      |
| 8        | 1032     | I          | 1996     | i          | 1935     |
| 9        | 952      | J          | 1271     | j          | 1366     |
|          |          | K          | 973      | k          | 717      |
|          |          | L          | 1988     | l          | 2074     |
|          |          | M          | 1812     | m          | 1253     |
|          |          | N          | 1942     | n          | 1534     |
|          |          | O          | 2046     | o          | 1336     |
|          |          | P          | 1681     | p          | 1382     |
|          |          | Q          | 1036     | q          | 734      |
|          |          | R          | 1925     | r          | 1671     |
|          |          | S          | 1927     | s          | 1626     |
|          |          | T          | 1587     | t          | 1669     |
|          |          | U          | 1927     | u          | 1530     |
|          |          | V          | 1443     | v          | 1282     |
|          |          | X          | 796      | x          | 712      |
|          |          | Y          | 360      | y          | 342      |
|          |          | Z          | 1554     | z          | 1307     |

**Tabla 5-1: Cantidad de muestras por caracter de test para cada escenario del Experimento 1: indubitados vs indubitados.**

## 5.1.2 Protocolo Experimento 2

Para el Experimento 2 se utilizaron las imágenes de los indubitados como entrenamiento y las imágenes de los dubitados como test.

Con el fin de tener resultados comparables, se mantuvo el mismo conjunto de entrenamiento.

En este experimento únicamente cambian las muestras de test, que ahora son todas las imágenes de caracteres del conjunto de dubitados. En la Tabla 5-2 se indica la cantidad de muestras para cada escenario del Experimento 2, en el conjunto de test. Como puede observarse, este es menor que en el Experimento 1, y además, de peor calidad en numerosas ocasiones.

| NÚMEROS  |          | MAYÚSCULAS |          | MINÚSCULAS |          |
|----------|----------|------------|----------|------------|----------|
| Caracter | Muestras | Caracter   | Muestras | Caracter   | Muestras |
| 0        | 145      | A          | 214      | a          | 184      |
| 1        | 163      | B          | 164      | b          | 145      |
| 2        | 142      | C          | 143      | c          | 131      |
| 3        | 130      | D          | 155      | d          | 161      |
| 4        | 118      | E          | 167      | e          | 172      |
| 5        | 115      | F          | 88       | f          | 92       |
| 6        | 124      | G          | 142      | g          | 135      |
| 7        | 105      | H          | 122      | h          | 125      |
| 8        | 112      | I          | 165      | i          | 164      |
| 9        | 104      | J          | 114      | j          | 100      |
|          |          | K          | 119      | k          | 120      |
|          |          | L          | 157      | l          | 169      |
|          |          | M          | 156      | m          | 132      |
|          |          | N          | 174      | n          | 169      |
|          |          | O          | 165      | o          | 171      |
|          |          | P          | 162      | p          | 140      |
|          |          | Q          | 52       | q          | 79       |
|          |          | R          | 170      | r          | 170      |
|          |          | S          | 161      | s          | 170      |
|          |          | T          | 145      | t          | 181      |
|          |          | U          | 145      | u          | 152      |
|          |          | V          | 107      | v          | 94       |
|          |          | X          | 70       | x          | 101      |
|          |          | Y          | 54       | y          | 76       |
|          |          | Z          | 131      | z          | 133      |

**Tabla 5-2: Cantidad de muestras por caracter de test para cada escenario del Experimento 2: dubitados vs indubitados.**

## **5.2 Sistema de reconocimiento de caracteres**

El sistema de reconocimiento de caracteres utilizado para la obtención de resultados de referencia de la base de datos adquirida es un software que se ejecuta con Matlab, llamado ScatNet (v0.2) [32]. El software no requiere ninguna instalación: sólo hay que añadir las carpetas con los correspondientes ficheros en el directorio de la herramienta Matlab para comenzar a utilizarlo. Las carpetas con el contenido de los ficheros se pueden descargar gratuitamente [32].

El sistema de reconocimiento soporta las características de las bases de datos: MNIST, USPS, Caltech, Brodatz, CURE y Yaleb. Como nuestra base de datos tiene las mismas características que MNIST, se utilizaros los ficheros correspondientes a ésta. Sin embargo, con estos ficheros, el software sólo era capaz de reconocer caracteres de tipo numérico. Para la obtención de resultados de nuestra base de datos, hubo que adaptar dicho ficheros para que el software también reconociese letras mayúsculas y minúsculas.

El sistema nos muestra como salida la tasa de error del conjunto de test que se introduce. Para ello, el sistema obtiene las características de cada tipo de caracter del conjunto de entrenamiento y almacena los modelos creados. Los tipos de caracteres alfanuméricos son:

- Tipos de números: del “0” al “9”.
- Tipos de letras mayúsculas: de la “A” a la “Z”.
- Tipos de letras minúsculas: de la “a” a la “z”.

De tal forma que cada tipo queda definido en un vector de características. Para el cálculo de estos vectores el sistema de reconocimiento implementa operadores *Scattering* de segunda dimensión [31]. Un vector *Scattering* es un descriptor local que incluye información multi-escala y multi-dirección, utiliza descomposición de cascadas de *wavelets* y módulo complejo. Incluye coeficientes similares a descriptores *SIFT* junto a coeficientes de múltiples escalas y orientaciones.

Siendo  $R_{\gamma}x$  la rotación de un ángulo  $\gamma$  de un vector bidimensional  $x \in \mathbb{R}^2$ , los *wavelets* direccionales son obtenidos a través de rotar un solo *wavelet*  $\psi$ , a lo largo de  $K$  ángulos  $\gamma$ . Escalándolos con  $2^j$  campos:

$$\psi_{j,\gamma}(x) = 2^{-2j} \psi(2^{-j}R_{\gamma}x).$$

La transformada *wavelet* direccional de  $f$  en una posición  $x$  para escalas  $2^j < 2^J$  es un vector de coeficientes:

$$W_j f(x) = (f * \psi_{j,\gamma}(x), f * \Phi_j(x))_{j < J, \gamma} \quad (1)$$

Donde  $\Phi_j(x) = 2^{-2J} \Phi(2^{-J}x)$  es un filtro paso bajo que contiene bajas frecuencias de  $f$  sobre la escala  $2^J$ .

Los vectores *Scattering* empleados definen un caracter manuscrito independientemente de su escala y dirección, siendo además robusto frente a pequeñas deformaciones.

Existen vectores descriptivos que también consiguen invariancia en la orientación del caracter, pero tienen una desventaja: reducen la información del caracter. Sin embargo, con las operaciones de *Scattering* se recuperan las altas frecuencias perdidas y las transforma en coeficientes de co-ocurrencia de múltiples escalas y orientaciones. Los vectores de *Scattering* son capaces de proporcionar información precisa de estructuras complejas tales como esquinas del caracter y cruces.

Una vez obtenidos los vectores *Scattering* del conjunto de entrenamiento (vectores de  $N$  dimensiones), se aplica un clasificador PCA [31]. Dicho clasificador reduce a  $d < N$  dimensiones, reteniendo los autovectores de mayor varianza. Los nuevos vectores de  $d$  dimensiones son representados espacialmente formando la base PCA, con la que se dividirá el espacio en tantas regiones como tipos de caracteres tenga el conjunto de entrenamiento. En base a las regiones creadas se hace la clasificación de los caracteres del conjunto de test, de tal manera que con cada muestra se obtiene una lista de *scores* según la distancia a la que esté de cada región PCA (identificando la muestra con el tipo de carácter de la región en la que ha sido representada, la clase con mayor *score*).

Por último, el sistema comprueba si su decisión ha sido acertada, ya que según la organización de las matrices de vectores donde se almacena el conjunto de test, se puede

saber el valor real de cada imagen. De esta manera calcula la tasa de error que tiene frente al conjunto de test introducido.

De este modo, el reconocedor de caracteres lleva a cabo los siguientes pasos:

- i. Se calcula el vector *Scattering* de características de cada tipo de caracter del conjunto de entrenamiento.
- ii. Cada vector obtenido es representado espacialmente mediante un clasificador PCA. Se obtiene así el mapa espacial con todos los tipos de caracteres del conjunto de entrenamiento localizados.
- iii. A continuación, se calcula el vector de características (con operadores *Scattering*) de cada imagen del conjunto de test y mediante el clasificador PCA se sitúa cada caracter en el mapa espacial previamente creado a partir del conjunto de entrenamiento.
- iv. Posteriormente, se selecciona la mejor aproximación espacial (el mejor *score*) para cada muestra del conjunto de test.
- v. Por último se calcula la tasa de error cometida.

En la Figura 5-2 se muestra un ejemplo del funcionamiento de un clasificador PCA bidimensional. En este ejemplo hay 4 clases de señales, cada una de ellas se sitúa en el espacio geométrico que le corresponde. Tras localizar espacialmente cada muestra del conjunto de entrenamiento, el clasificador crea 4 zonas. A la hora de evaluar las muestras del conjunto de test, se asociará a cada una de ellas una clase dependiendo en la zona que estén localizadas.

En el software empleado se pueden modificar el valor de los parámetros que se van a utilizar para evaluar la base de datos, con el objetivo de adaptarlos al problema en cuestión: reconocimiento de caracteres manuscritos. Estos parámetros están directamente relacionados con los operadores de *Scattering*, siendo:

- Parámetro J: la escala máxima.
- Parámetro L: número máximo de orientaciones.
- Parámetro M: el orden máximo de los coeficientes.

La transformada *wavelet* (1) muestra que las altas frecuencias eliminadas en  $|f * \psi_{j_1, r_1}| * \Phi_J$  por la convolución con  $\Phi_J$  son recuperadas por las convoluciones con *wavelets*  $|f * \psi_{j_1, r_1}| * \psi_{j_2, r_2}$  en escalas  $2^{j_2} < 2^J$ . Para ser insensible a las traslaciones locales y reducir la variabilidad de estos coeficientes, su fase compleja es eliminada con el módulo, y promediada por  $\Phi_J$ :

$$\| |f * \psi_{j_1, r_1}| * \psi_{j_2, r_2} \| * \Phi_J$$

Se denominan coeficientes de *Scattering* por ser el resultado de los cruces de  $f$  con los dos *wavelets*. Dan información de la escala ( $2^{j_1}, 2^{j_2}$ , etc.) y la orientación ( $r_1, r_2$ , etc.). Además el vector *Scattering* agrega a estos coeficientes un orden máximo ( $m$ ).

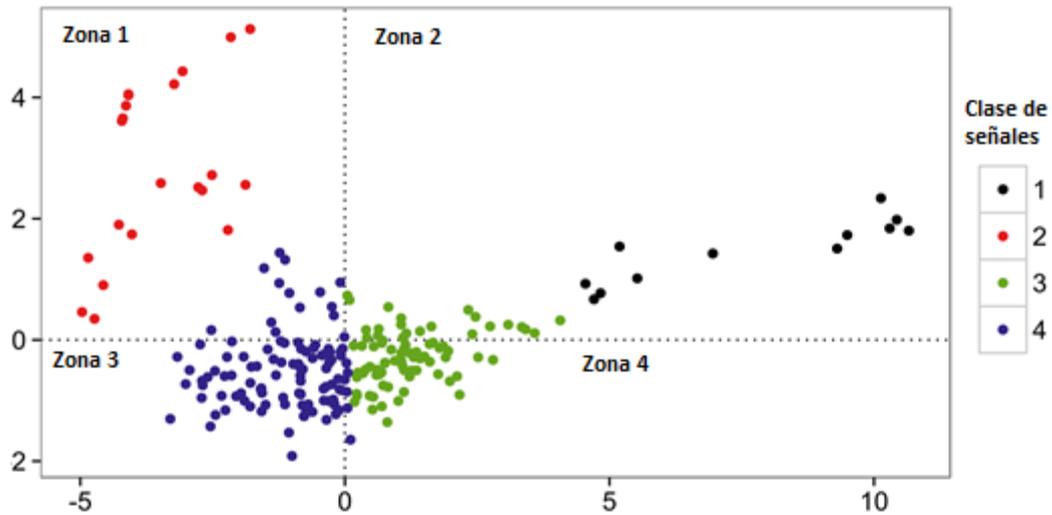


Figura 5-2: Ejemplo en pequeña escala de clasificador PCA

### 5.3 Resultados de referencia

Los resultados de referencia se obtuvieron con el software explicado en la sección anterior (apartado 5.2).

Previo a la obtención de resultados, se buscaron los mejores valores para los parámetros del sistema de reconocimiento de tal forma que estos fuesen lo más cercanos posibles al óptimo.

Se realizaron cuatro pruebas para optimizar los parámetros de cada uno de los distintos escenarios (Números, Mayúsculas, Minúsculas, y Todos), probando todas las combinaciones posibles de parámetros en los rangos  $J = [2,6]$ ,  $M = [2,6]$ ,  $L = [2,8]$  (los autores de [31] recomendaban  $J=3$ ,  $M=2$ )

|                    |            | Parámetros |   |   |
|--------------------|------------|------------|---|---|
|                    |            | J          | M | L |
| Tipo de evaluación | Números    | 4          | 4 | 5 |
|                    | Mayúsculas | 4          | 4 | 8 |
|                    | Minúsculas | 4          | 3 | 5 |
|                    | Todos      | 4          | 4 | 8 |

Taba 5-3: Valor de los parámetros para cada modalidad de evaluación

Tras obtener los mejores parámetros para cada escenario (tabla 5-3), se procedió a la obtención de resultados para cada experimento y escenario:

- Experimento 1: indubitados vs indubitados. Con 4 modalidades: Números, mayúsculas, minúsculas, todo.
- Experimento 2: indubitados vs dubitados. Con 4 modalidades: Números, mayúsculas, minúsculas, todo.

En las próximas secciones (5.3.1 y 5.3.2) se muestran los resultados obtenidos. Para concluir, en la sección 5.3.3, se analizan y comparan los resultados de ambos experimentos.

### 5.3.1 Resultados Experimento 1: indubitados vs indubitados

Las tasas de error obtenidas en el Experimento 1 se muestran en la tabla 5-4

| Tasas de error Experimento 1: indubitados vs indubitados |            |            |        |
|--|------------|------------|--------|
| Números  | Mayúsculas | Minúsculas | Todo   |
| 2,6%   | 10,49%     | 26,5%      | 30,86% |

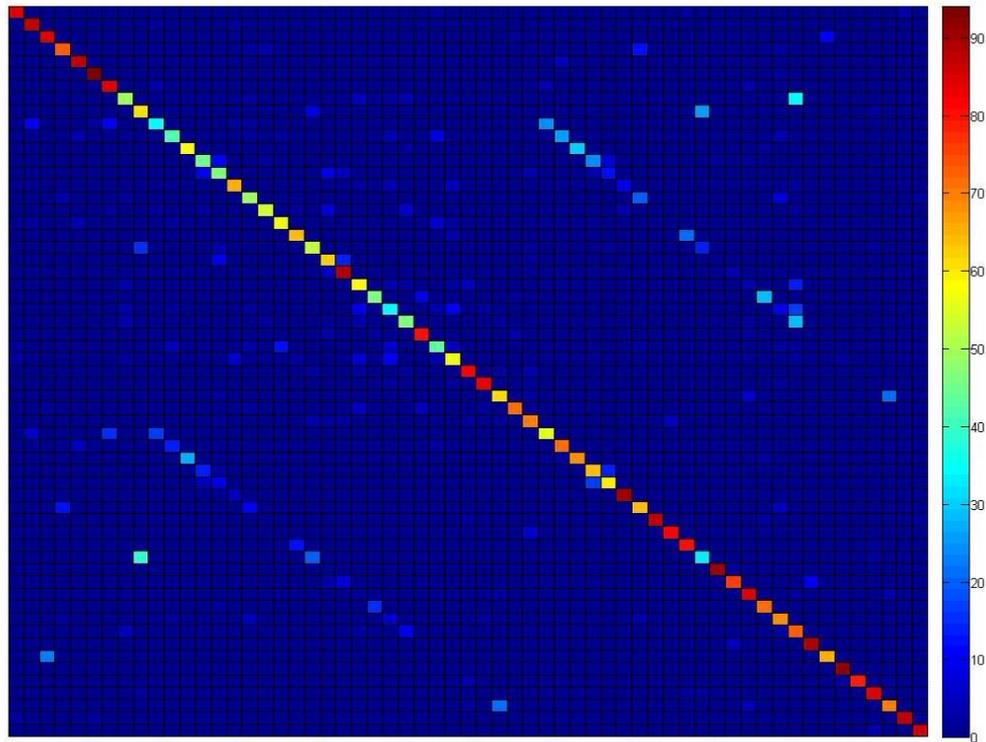
**Tabla 5-4: Tasas de error de los 4 escenarios del Experimento 1**

A continuación, en la tabla 5-5, se muestra la matriz de confusión del escenario Todo (donde se comparan números, letras mayúsculas y minúsculas). En la primera columna de la izquierda está el tipo de carácter que se ha evaluado y en la primera fila las posibles clases de caracteres que tiene el sistema de reconocimiento (obtenido del conjunto de entrenamiento). De tal forma que en cada celda de la matriz, se muestra en tanto por ciento las veces que el carácter de la columna de la izquierda ha sido aproximado a cada tipo de carácter posible. Las celdas que se encuentran en la diagonal de la matriz de confusión son los cruces de los caracteres evaluados con su correspondiente tipo de carácter, es por ello que en la diagonal, por lo general, se encuentran los valores más altos.

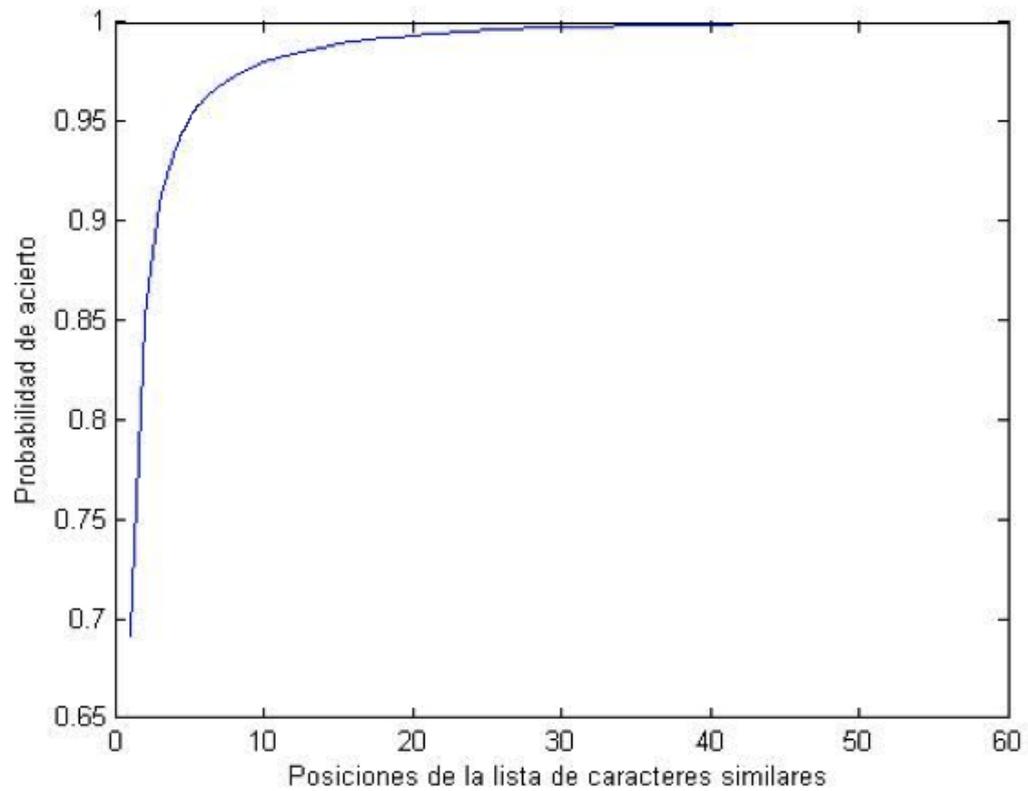
En la figura 5-3, se muestran una ilustración visual de la matriz de confusión previamente explicada. En esta figura se puede apreciar claramente los valores más elevados en la diagonal.

Por último se obtuvo la curva CMC del escenario Todo. Cuando el sistema de reconocimiento asigna una clase de muestra a un carácter del conjunto de test, elabora un rango de clases para la imagen con los tipos de muestra posible, asignando a la imagen de test el tipo de clase al que más se aproxima. Nos pareció interesante elaborar la curva CMC para ver que probabilidad de acierto hay en que el tipo de carácter acertado esté entre los  $n$  primeros puestos del rango, con  $n$  entre 1 y 60 (número total de clases). A continuación, en la figura 5-4, se muestra la curva CMC para el escenario Todo del Experimento 1.





**Figura 5-3: Matriz de confusión del escenario Todo del Experimento 1 representada con colores**



**Figura 5-4: Curva CMC del escenario Todo del Experimento 1. En la posición 5 de la lista, hay una probabilidad de acierto del 95.18%**

### 5.3.2 Resultados Experimento 2: dubitados vs indubitados

Las tasas de error obtenidas en el Experimento 2 se muestran en la tabla 5-6. A continuación, en la tabla 5-7 se ilustra la matriz de confusión para el escenario Todo, donde se pueden observar características muy similares a las comentadas para la matriz de confusión del Experimento 1. En la ilustración visual de la misma, figura 5-5, también se puede apreciar los valores elevados en la diagonal.

Por último, en la figura 5-6 se muestra la curva CMC para el escenario Todo del Experimento 2.

| Tasas de error Experimento 2: dubitados vs indubitados |            |            |       |
|--|------------|------------|-------|
| Números  | Mayúsculas | Minúsculas | Todo  |
| 6,52   | 12,52%     | 27,44%     | 36,9% |

Tabla 5-6: Tasas de error de los 4 escenarios del Experimento 2

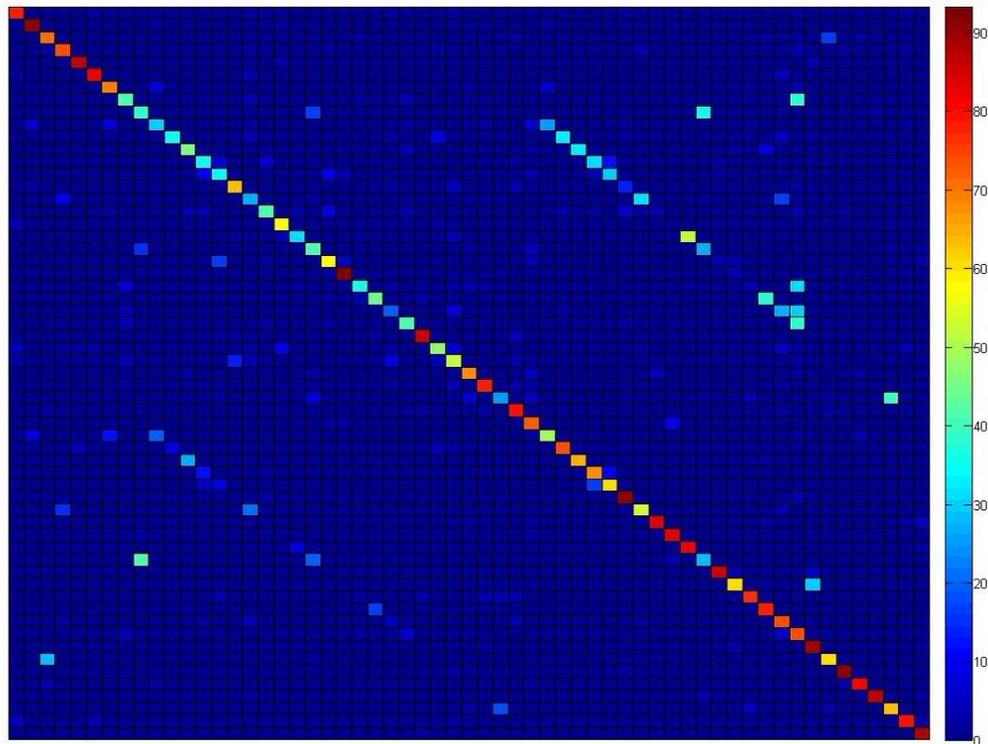
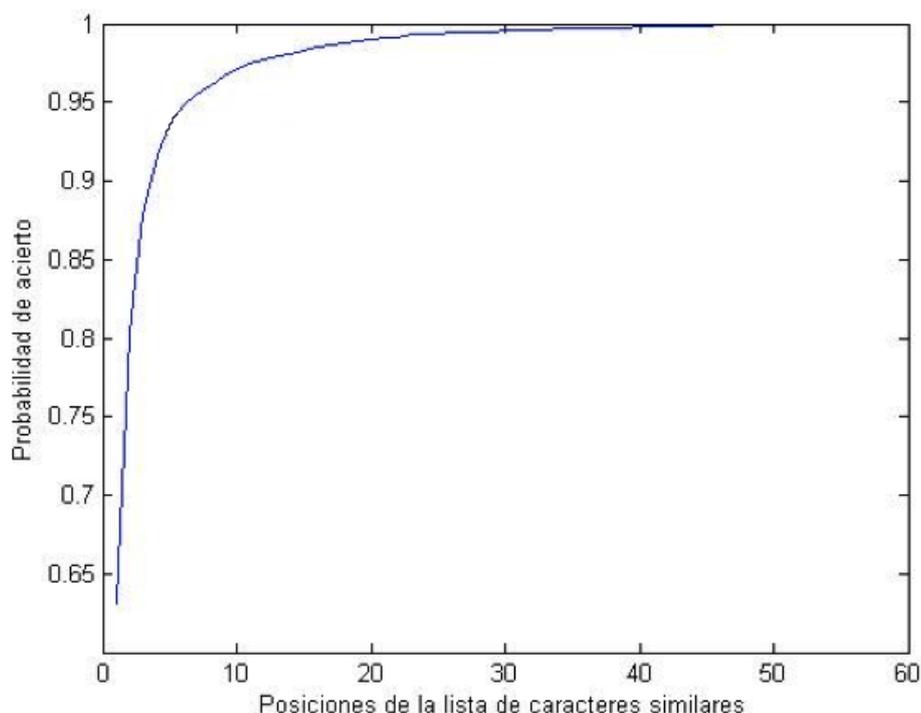


Figura 5-5: Matriz de confusión del escenario Todo del Experimento 2 representada con colores





**Figura 5-6: Curva CMC de la evaluación “Todo” del Experimento 2. En la posición 5 de la lista, hay una probabilidad de acierto del 93.61%**

### 5.3.3 Discusión

En primer lugar podemos observar que para cada escenario la tasa de error es mayor en el Experimento 2 que en el Experimento 1. Como se ha explicado previamente (secciones 5.1.1. y 5.1.2), el conjunto de entrenamiento de ambos experimentos es el mismo, lo que cambia es el conjunto de test. En el Experimento 1 el conjunto de test son caracteres de indubitados, al igual que en el conjunto de entrenamiento, mientras que en el Experimento 2, su conjunto de test son caracteres de dubitados. Estos últimos son imágenes de caracteres recogidas en ámbitos forenses por lo que los utensilios de escritura con los que están hechas estas muestras son variables en cada documento y no tienen nada en común con los documentos del conjunto de entrenamiento. Sin embargo en el Experimento 1 tanto el conjunto de entrenamiento como el conjunto de test son caracteres elaborados con los mismos utensilios de escritura y bajo supervisión. Es razonable que por ello, las tasas de error en los escenarios del Experimento 1 sean menores que las tasas de error para el Experimento 2.

Se puede apreciar asimismo que la menor tasa de error es para el escenario de Números (donde se comparan números) y la tasa de error más alta se obtiene en el escenario Todo (donde se comparan números, letras mayúsculas y minúsculas). Esto tiene una explicación matemática y es que en la evaluación de números sólo hay 10 tipos de caracteres y en la evaluación de todos los caracteres alfanuméricos hay 60 tipos de caracteres (10 números, 25 letras mayúsculas y 25 minúsculas), por lo tanto la probabilidad de fallo es menor en la evaluación de números que en la evaluación de todo. La probabilidad de fallar según la cantidad de tipos de caracteres que se evalúen es:

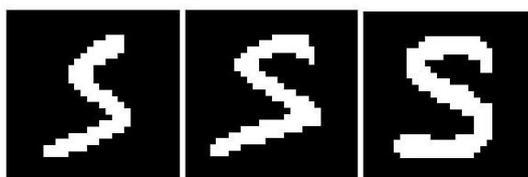
$$\text{Probabilidad (fallar)} = 1 - (1/\text{Numero de clases})$$

En la tabla 5-8 se muestra la probabilidad de fallar según el escenario.

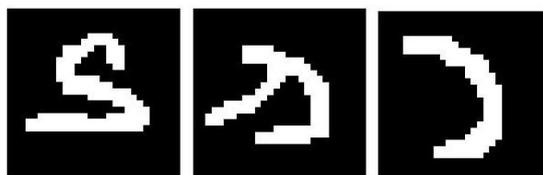
|           |            | Probabilidad de fallar |
|-----------|------------|------------------------|
| Escenario | Números    | 1- (1/10) → 90%        |
|           | Mayúsculas | 1- (1/25) → 96%        |
|           | Minúsculas | 1- (1/25) → 96%        |
|           | Todo       | 1- (1/60) → 98.3%      |

**Tabla 5-8: Probabilidad de fallo según escenario**

A priori se podría decir que la tasa de error de los escenarios Mayúsculas y Minúsculas debería ser parecida según las explicaciones matemáticas, pero como se puede observar en las tablas 5-4 y 5-6 esto no sucede así. La tasa de error en los escenarios Minúsculas es más del doble de la tasa de error de Mayúsculas, observando cada experimento por separado. Esto es debido a que los caracteres manuscritos de letras minúsculas tienen mucha mayor variabilidad de formas para representar un único carácter que las letras mayúsculas. Por ejemplo la letra “s” en minúsculas tiene varias formas para su representación (figura 5-7), mientras que su equivalente en mayúsculas “S” sólo tiene una forma de representarse, aunque varíe la curvatura del carácter según el escritor (figura 5-8).



**Figura 5-7: Letras “S” mayúsculas de la base de datos procesada**



**Figura 5-8: Letras “s” minúsculas de la base de datos procesada**

Por lo tanto las tasas de error en los escenarios Minúsculas es mayor que para Mayúsculas, debido a que la representación de las letras mayúsculas es mucho más uniforme y precisa que la representación de las letras minúsculas, dando lugar a menos casos de equivocación al sistema de reconocimiento utilizado y por lo tanto, menor tasa de error.

Cabe destacar que el resultado obtenido para la tasa de error de ambos experimentos en el escenario de Mayúsculas (10,49% en el Experimento 1 y 12,52% en el Experimento 2) es mayor que MNIST pero sus resultados son sólo con números. Si comparamos con Firemaker que sus experimentos son sólo con mayúsculas (entre 17,79% y 32,99% de tasa de error) las tasas de error obtenidas en estos resultados de referencia son coherentes.

Otros tipos de casos de equivocación pueden ser entre caracteres de distintas modalidades alfanuméricas, como se muestra en la tabla 5-9. Para observar este tipo de fallos se obtuvieron las matrices de confusión presentadas en las secciones anteriores.

En la Tabla 5-5 se puede observar, en amarillo, un ejemplo de un caso en el cual el valor más alto de semejanza no se encuentra en su correspondiente clase, es el caso de la evaluación de la letra “O”, el 33% de los casos el sistema de reconocimiento lo aproxima a una “O” y en el 39% de los casos a un “0”. Al igual que en la matriz del Experimento 2 el 28% de los casos el sistema de reconocimiento aproxima la “O” a su correspondiente valor y en el 41% de los casos lo aproxima a un “0”. Otro ejemplo en el Experimento 2, tabla 5-7, es la evaluación del número “1” donde el valor más alto sí se presenta en el acertado, el 41% de veces lo aproxima a “1”, pero el 39% de las veces, el sistema de reconocimiento lo aproxima a una “I”. Como puede observarse en los ejemplos de la tabla 5-9, un humano confundiría fácilmente dichos caracteres

| Caracter a evaluar |          | Clase asignada |
|--------------------|----------|----------------|
| Imagen             | Caracter |                |
|                    | G        | 6              |
|                    | H        | M              |
|                    | M        | H              |
|                    | O        | 0              |
|                    | R        | A              |
|                    | S        | 5              |
|                    | a        | Q              |
|                    | f        | t              |
|                    | g        | 9              |
|                    | n        | u              |
|                    | q        | 9              |
|                    | s        | J              |
|                    | 1        | I              |
|                    | 2        | Z              |
|                    | 8        | B              |

**Tabla 5-9: Ejemplos de asignación de clases erróneas de caracteres del conjunto de test del escenario Todo**

En las figuras ilustrativas de las matrices de confusión (figuras 5-3 y 5-5), se ve claramente que la diagonal contiene los colores más cálidos, correspondientes a un valor alto de porcentaje. Además es curioso cómo aparecen otras dos diagonales con colores no tan fríos como el resto de la matriz pero tampoco tan cálidos como la diagonal. Estas diagonales, si nos fijamos en las tablas 5-5 y 5-7, corresponden al cruce de los caracteres de letras mayúsculas con los caracteres de las letras minúsculas. Este efecto es debido a que la representación de muchas letras mayúsculas es muy similar a la correspondiente en letras minúsculas sólo diferenciándose en el tamaño, y varias veces el sistema de reconocimiento no puede distinguirlos, ya que al pre-procesar las imágenes se perdió la información relativa al tamaño original del carácter. Por esto aparecen las otras dos diagonales. Estos fallos de nuevo podrían ser cometidos también por un humano.

Tras comentar y analizar las matrices de confusión, se muestra en la tabla 5-9 algunos casos de asignación de clases erróneas en los que el sistema de reconocimiento asigna al carácter del conjunto de test otro tipo de carácter que no le corresponde. En algunos casos estos errores que comete el sistema de reconocimiento podrían ser errores humanos, como en el ejemplo de la evaluación de la “q” que la aproxima a un “9”, a ojos humanos se puede reconocer como un “9” también.

Por último vamos a analizar las curvas CMC de ambos experimentos. Para el escenario Todo en el Experimento 1 se obtuvo una tasa de error de 30,86% (Tabla 5-4), en la curva CMC, figura 5-4, se puede observar que hay una probabilidad de acierto del 95.18% de que el carácter correcto esté entre los 5 primeros del rango. Por lo tanto puede resultar que el sistema de reconocimiento no acierte a la primera con el carácter adecuado (esto sólo lo hace en un 30.86% de las veces) pero cabe destacar que el carácter correcto se encontrará en un 95.18% de los casos entre las 5 primeras clases propuestas.

Análogamente, para el escenario Todo en el Experimento 2 se obtuvo una tasa de error de 36,9% (Tabla 5-7) y en esta curva CMC, figura 5-6, hay una probabilidad de acierto del 93.61% de que el carácter acertado esté entre los 5 primeros.



## 6 Conclusiones y trabajo futuro

---

En el presente proyecto se ha adquirido y procesado una base de datos de caracteres manuscritos compuesta por escritores españoles. La base de datos contiene caracteres manuscritos de tres tipos: números, letras mayúsculas y letras minúsculas. Como hemos visto a lo largo de este proyecto, es una base de datos muy amplia con más variedad que la referente hoy en día, MNIST [1].

En primer lugar se ha adquirido una base de datos formada por caracteres manuscritos de 527 personas de las cuales 478 son personas identificadas (indubitados) y 49 son personas sin identificar (dubitados). Para la segmentación manual de los documentos se utilizó la herramienta Biógrafo descrita en este documento.

Para la obtención de resultados de referencia se procesaron las imágenes de la base de datos adquirida. Las características finales que tienen las imágenes de caracteres son las mismas que tienen las de MNIST, de tal forma que las características de las imágenes diesen lugar a una fácil futura incorporación de nuestra base de datos en sistemas de reconocimiento de caracteres manuscritos.

Por lo tanto, la base de datos adquirida en este proyecto está disponible en dos formatos; imágenes del tipo “.bmp” sin procesar, e imágenes procesadas con las características descritas en este documento para su directa aplicación en sistemas de reconocimiento ya adaptados a MNIST.

Por último, se obtuvieron resultados de la base de datos procesada con un sistema de reconocimiento de caracteres [32] descrito en este documento que utiliza vectores de *Scattering* y un clasificador PCA [31]. Dicho sistema de reconocimiento ya había sido usado sobre MNIST. Las tasas de error obtenidas con la base de datos servirán como referencia a mejorar para la comunidad de investigación de reconocimiento de caracteres manuscritos.

El proyecto abre una amplia gama de posibilidades para el desarrollo y mejora de sistemas de reconocimiento de caracteres manuscritos. La principal contribución de este proyecto es una amplia base de datos en dos formatos:

- Imágenes del tipo “.bmp” sin procesar.
- Imágenes procesadas con las características de MNIST.

Como líneas de trabajo futuro, este proyecto aporta:

- Los resultados de referencia, obtenidos con el sistema de reconocimiento de caracteres, como un punto de partida a partir del cual se puede continuar la labor de investigación con la obtención de más resultados de referencia para diferentes algoritmos, métodos y clasificadores de caracteres manuscritos.



## Referencias

---

- [1] Yann LeCun, Corina Cortes, Christopher J.C. Burgues, “The MNIST database of handwritten digits”, <http://yann.lecun.com/exdb/mnist/>
- [2] S. Impedovo, “Frontiers in Handwriting Recognition”, en “Fundamentals in Handwriting Recognition”, NATO-ASI Series, Springer- Verlag Publ., Berlin, 1994
- [3] Réjean Plamondon, Fellow, IEEE and Sargur N. Srihari, Fellow, IEEE. “On-Line and Off-Line Handwritten Recognition: A Comprehensive survey” IEEE Transactions on Pattern analysis and machine intelligence, Vol 22, No. 1, January 2000.
- [4] M. Cheriet and C.Y. Sue. “Extraction of Key Letters Script Recognition Pattern Recognition Letters”, páginas 1009-1017, 1993
- [5] Universidad de Oviendo “Aplicaciones de la optimización. Algoritmos para la eliminación de ruido”  
[http://www.unioviendo.es/compnum/laboratorios\\_web/laborat08\\_optimizacion1/laborat08\\_optimizacion1.html#11](http://www.unioviendo.es/compnum/laboratorios_web/laborat08_optimizacion1/laborat08_optimizacion1.html#11)
- [6] R. Santiago Montero, J. M. López Márquez, M. Ornelas Rodríguez “Aplicando filtros de Gabor a segmentación de caracteres alfanuméricos en placas vehiculares”, Mexico, 2011
- [7] José R. Hilera González, Juan P. Romero Villaverde, José A. Gutiérrez de Mesa, “Sistema de reconocimiento óptico de caracteres (OCR) con redes neuronales” 1996
- [8] O.D. Trier, A.K. Jain, T. Taxt. “Feature extraction methods for character recognition— a survey, Pattern Recognition” páginas 641–662, 1996
- [9] F. Kimura, et al., “Evaluation and synthesis of feature vectors for handwritten numeral recognition, IEICE Trans. Inform. Systems” páginas 436–442, 1996
- [10] M. Shi, Y. Fujisawa, T. Wakabayashi, F. Kimura, “Handwritten numeral recognition using gradient and curvature of gray scale image, Pattern Recognition” páginas 2051–2059, 2002
- [11] S. Lee, S.N. Srihari, “Handprinted digit recognition: a comparison of algorithms, in: Proceedings of the Third International Workshop on Frontiers of Handwriting Recognition”, páginas 153–164, Buffalo, NY, 1993
- [12] L. Liu, Y.-J. Liu, R.-W. Dai, “Preprocessing and statistical/structural feature extraction for handwritten numeral recognition” páginas 161–168, Singapore, 1997
- [13] L. Heutte, T. Paquet, J.V. Moreau, Y. Lecourtier, C. Olivier, “A structural/statistical feature based vector for handwritten character recognition, Pattern Recognition Lett.” páginas 629–641, 1998
- [14] K. Fukunaga, “Introduction to Statistical Pattern Recognition”, 2nd Edition, Academic Press, New York, 1990.
- [15] R.O. Duda, P.E. Hart, D.G. Stork, “Pattern Classification”, 2nd Edition, Wiley Interscience, New York, 2000.
- [16] A.K. Jain, R.P.W. Duin, J. Mao, “Statistical pattern recognition: a review, IEEE Trans. Pattern Anal. Mach. Intell.” páginas 4–37, 2000
- [17] C.M. Bishop, “Neural Networks for Pattern Recognition, Clarendon Press”, Oxford, 1995.
- [18] L. Holmström, et al., “Neural and statistical classifiers— taxonomy and two case studies”, IEEE Trans. Neural Networks, páginas 5–17, 1997
- [19] F. Kimura, K. Takashina, S. Tsuruoka, Y. Miyake, “Modified quadratic discriminant functions and the application to Chinese character recognition”, IEEE Trans. Pattern Anal. Mach. Intell. páginas 149–153, 1987

- [20] F. Kimura, M. Shridhar, “Handwritten numeral recognition based on multiple algorithms”, Pattern Recognition, páginas 969–981, 1991
- [21] D.E. Rumelhart, G.E. Hinton, R.J. Williams, “Learning representations by back-propagation errors”, Nature páginas 533–536, 1986
- [22] U. Krebel, J. Schuürmann, “Pattern classification techniques based on function approximation”, H. Bunke, P.S.P. Wang (Eds.), Handbook of Character Recognition and Document Image Analysis, World Scientific, páginas 49–78, Singapore, 1997
- [23] C.J.C. Burges, “A tutorial on support vector machines for pattern recognition, Knowledge Discovery Data Mining”, páginas 1–43, 1998
- [24] V. Vapnik, “The Nature of Statistical Learning Theory”, Springer, New Work, 1995.
- [25] Laurens Van der Maaten, “A New Benchmark Dataset for Handwritten Character Recognition”, Tilburg University, páginas 2-5, April 26, 2009.
- [26] Daniel Keisers “Experimental results on the USPS database”  
<http://www-i6.informatik.rwth-aachen.de/~keyisers/Pubs/SPR2002/node10.html> , 2002
- [27] Javier Galbally, Santiago Gonzalez-Dominguez, Julian Fierrez, and Javier Ortega-Garcia, “Biografo: An Integrated Tool for Forensic Writer Identification”, Biometric Recognition Group-ATVS, EPS, Universidad Autonoma de Madrid, 2012
- [28] MathWorks, Documentation Center  
<http://www.mathworks.es/es/help/images/ref/bwlabel.html>
- [29] MathWorks, Documentation Center  
<http://www.mathworks.es/es/help/matlab/ref/mode.html>
- [30] MathWorks, DocumentationCenter  
<http://www.mathworks.com/matlabcentral/fileexchange/5457>
- [31] Joan Bruna y Stéphane Mallat, “Classification with Scattering Operator”, Ecole Polytechnique, Palaiseau 2011
- [32] Software de reconocimiento de caracteres manuscritos.  
<http://www.di.ens.fr/data/software/> Ecole Normale Superieure, Paris. Versión Marzo de 2011

# PRESUPUESTO

## 1) Ejecución Material

- Compra de ordenador personal (Software incluido)..... 2.000 €
- Escaneo de los documentos escritos (persona contratada)..... 500 €
- Material de oficina ..... 150 €
- 2 discos duros de almacenamiento ..... 120 €
- Total de ejecución material ..... 2.770 €

## 2) Gastos generales

- sobre Ejecución Material ..... 443.2 €

## 3) Beneficio Industrial

- 6 % sobre Ejecución Material ..... 166.2 €

## 4) Honorarios Proyecto

- 1200 horas a 15 € / hora..... 18000 €

## 5) Material fungible

- Gastos de impresión..... 180 €
- Encuadernación..... 70 €

## 6) Subtotal del presupuesto

- Subtotal Presupuesto..... 21629.4 €

## 7) I.V.A. aplicable

- 21% Subtotal Presupuesto ..... 4542.2 €

## 8) Total presupuesto

- Total Presupuesto..... 26171.6 €

Madrid, Octubre de 2014

El Ingeniero Jefe de Proyecto

Fdo.: Sara García-Mina Martínez  
Ingeniero de Telecomunicación



## **PLIEGO DE CONDICIONES**

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, de la CONSTRUCCIÓN DE UNA NUEVA BASE DE DATOS PARA EL RECONOCIMIENTO AUTOMÁTICO DE CARACTERES MANUSCRITOS Y GENERACIÓN DE RESULTADOS DE REFERENCIA. En lo que sigue, se supondrá que el proyecto ha sido encargado por una empresa cliente a una empresa consultora con la finalidad de realizar dicho sistema. Dicha empresa ha debido desarrollar una línea de investigación con objeto de elaborar el proyecto. Esta línea de investigación, junto con el posterior desarrollo de los programas está amparada por las condiciones particulares del siguiente pliego.

Supuesto que la utilización industrial de los métodos recogidos en el presente proyecto ha sido decidida por parte de la empresa cliente o de otras, la obra a realizar se regulará por las siguientes:

### **Condiciones generales**

1. La modalidad de contratación será el concurso. La adjudicación se hará, por tanto, a la proposición más favorable sin atender exclusivamente al valor económico, dependiendo de las mayores garantías ofrecidas. La empresa que somete el proyecto a concurso se reserva el derecho a declararlo desierto.

2. El montaje y mecanización completa de los equipos que intervengan será realizado totalmente por la empresa licitadora.

3. En la oferta, se hará constar el precio total por el que se compromete a realizar la obra y el tanto por ciento de baja que supone este precio en relación con un importe límite si este se hubiera fijado.

4. La obra se realizará bajo la dirección técnica de un Ingeniero Superior de Telecomunicación, auxiliado por el número de Ingenieros Técnicos y Programadores que se estime preciso para el desarrollo de la misma.

5. Aparte del Ingeniero Director, el contratista tendrá derecho a contratar al resto del personal, pudiendo ceder esta prerrogativa a favor del Ingeniero Director, quien no estará obligado a aceptarla.

6. El contratista tiene derecho a sacar copias a su costa de los planos, pliego de condiciones y presupuestos. El Ingeniero autor del proyecto autorizará con su firma las copias solicitadas por el contratista después de confrontarlas.

7. Se abonará al contratista la obra que realmente ejecute con sujeción al proyecto que sirvió de base para la contratación, a las modificaciones autorizadas por la superioridad o a las órdenes que con arreglo a sus facultades le hayan comunicado por escrito al Ingeniero Director de obras siempre que dicha obra se haya ajustado a los preceptos de los pliegos de condiciones, con arreglo a los cuales, se harán las modificaciones y la valoración de las diversas unidades sin que el importe total pueda exceder de los presupuestos aprobados. Por consiguiente, el número de unidades que se consignan en el proyecto o en el presupuesto, no podrá servirle de fundamento para entablar reclamaciones de ninguna clase, salvo en los casos de rescisión.

8. Tanto en las certificaciones de obras como en la liquidación final, se abonarán los trabajos realizados por el contratista a los precios de ejecución material que figuran en el presupuesto para cada unidad de la obra.

9. Si excepcionalmente se hubiera ejecutado algún trabajo que no se ajustase a las condiciones de la contrata pero que sin embargo es admisible a juicio del Ingeniero Director de obras, se dará conocimiento a la Dirección, proponiendo a la vez la rebaja de precios que el Ingeniero estime justa y si la Dirección resolviera aceptar la obra, quedará el contratista obligado a conformarse con la rebaja acordada.

10. Cuando se juzgue necesario emplear materiales o ejecutar obras que no figuren en el presupuesto de la contrata, se evaluará su importe a los precios asignados a otras obras o materiales análogos si los hubiere y cuando no, se discutirán entre el Ingeniero Director y el contratista, sometiéndolos a la aprobación de la Dirección. Los nuevos precios convenidos por uno u otro procedimiento, se sujetarán siempre al establecido en el punto anterior.

11. Cuando el contratista, con autorización del Ingeniero Director de obras, emplee materiales de calidad más elevada o de mayores dimensiones de lo estipulado en el proyecto, o sustituya una clase de fabricación por otra que tenga asignado mayor precio o ejecute con mayores dimensiones cualquier otra parte de las obras, o en general, introduzca en ellas cualquier modificación que sea beneficiosa a juicio del Ingeniero Director de obras, no tendrá derecho sin embargo, sino a lo que le correspondería si hubiera realizado la obra con estricta sujeción a lo proyectado y contratado.

12. Las cantidades calculadas para obras accesorias, aunque figuren por partidaalzada en el presupuesto final (general), no serán abonadas sino a los precios de la contrata, según las condiciones de la misma y los proyectos particulares que para ellas se formen, o en su defecto, por lo que resulte de su medición final.

13. El contratista queda obligado a abonar al Ingeniero autor del proyecto y director de obras así como a los Ingenieros Técnicos, el importe de sus respectivos honorarios facultativos por formación del proyecto, dirección técnica y administración en su caso, con arreglo a las tarifas y honorarios vigentes.

14. Concluida la ejecución de la obra, será reconocida por el Ingeniero Director que a tal efecto designe la empresa.

15. La garantía definitiva será del 4% del presupuesto y la provisional del 2%.

16. La forma de pago será por certificaciones mensuales de la obra ejecutada, de acuerdo con los precios del presupuesto, deducida la baja si la hubiera.

17. La fecha de comienzo de las obras será a partir de los 15 días naturales del replanteo oficial de las mismas y la definitiva, al año de haber ejecutado la provisional, procediéndose si no existe reclamación alguna, a la reclamación de la fianza.

18. Si el contratista al efectuar el replanteo, observase algún error en el proyecto, deberá comunicarlo en el plazo de quince días al Ingeniero Director de obras, pues transcurrido ese plazo será responsable de la exactitud del proyecto.

19. El contratista está obligado a designar una persona responsable que se entenderá con el Ingeniero Director de obras, o con el delegado que éste designe, para todo relacionado con ella. Al ser el Ingeniero Director de obras el que interpreta el proyecto, el contratista deberá consultarle cualquier duda que surja en su realización.

20. Durante la realización de la obra, se girarán visitas de inspección por personal facultativo de la empresa cliente, para hacer las comprobaciones que se crean oportunas. Es obligación del contratista, la conservación de la obra ya ejecutada hasta la recepción de la misma, por lo que el deterioro parcial o total de ella, aunque sea por agentes atmosféricos u otras causas, deberá ser reparado o reconstruido por su cuenta.

21. El contratista, deberá realizar la obra en el plazo mencionado a partir de la fecha del contrato, incurriendo en multa, por retraso de la ejecución siempre que éste no sea debido a causas de fuerza mayor. A la terminación de la obra, se hará una recepción provisional previo reconocimiento y examen por la dirección técnica, el depositario de efectos, el interventor y el jefe de servicio o un representante, estampando su conformidad el contratista.

22. Hecha la recepción provisional, se certificará al contratista el resto de la obra, reservándose la administración el importe de los gastos de conservación de la misma hasta su recepción definitiva y la fianza durante el tiempo señalado como plazo de garantía. La recepción definitiva se hará en las mismas condiciones que la provisional, extendiéndose el acta correspondiente. El Director Técnico propondrá a la Junta Económica la devolución de la fianza al contratista de acuerdo con las condiciones económicas legales establecidas.

23. Las tarifas para la determinación de honorarios, reguladas por orden de la Presidencia del Gobierno el 19 de Octubre de 1961, se aplicarán sobre el denominado en la actualidad "Presupuesto de Ejecución de Contrata" y anteriormente llamado "Presupuesto de Ejecución Material" que hoy designa otro concepto.

### **Condiciones particulares**

La empresa consultora, que ha desarrollado el presente proyecto, lo entregará a la empresa cliente bajo las condiciones generales ya formuladas, debiendo añadirse las siguientes condiciones particulares:

1. La propiedad intelectual de los procesos descritos y analizados en el presente trabajo, pertenece por entero a la empresa consultora representada por el Ingeniero Director del Proyecto.

2. La empresa consultora se reserva el derecho a la utilización total o parcial de los resultados de la investigación realizada para desarrollar el siguiente proyecto, bien para su publicación o bien para su uso en trabajos o proyectos posteriores, para la misma empresa cliente o para otra.

3. Cualquier tipo de reproducción aparte de las reseñadas en las condiciones generales, bien sea para uso particular de la empresa cliente, o para cualquier otra aplicación, contará con autorización expresa y por escrito del Ingeniero Director del Proyecto, que actuará en representación de la empresa consultora.

4. En la autorización se ha de hacer constar la aplicación a que se destinan sus reproducciones así como su cantidad.

5. En todas las reproducciones se indicará su procedencia, explicitando el nombre del proyecto, nombre del Ingeniero Director y de la empresa consultora.

6. Si el proyecto pasa la etapa de desarrollo, cualquier modificación que se realice sobre él, deberá ser notificada al Ingeniero Director del Proyecto y a criterio de éste, la empresa consultora decidirá aceptar o no la modificación propuesta.

7. Si la modificación se acepta, la empresa consultora se hará responsable al mismo nivel que el proyecto inicial del que resulta el añadirla.

8. Si la modificación no es aceptada, por el contrario, la empresa consultora declinará toda responsabilidad que se derive de la aplicación o influencia de la misma.

9. Si la empresa cliente decide desarrollar industrialmente uno o varios productos en los que resulte parcial o totalmente aplicable el estudio de este proyecto, deberá comunicarlo a la empresa consultora.

10. La empresa consultora no se responsabiliza de los efectos laterales que se puedan producir en el momento en que se utilice la herramienta objeto del presente proyecto para la realización de otras aplicaciones.

11. La empresa consultora tendrá prioridad respecto a otras en la elaboración de los proyectos auxiliares que fuese necesario desarrollar para dicha aplicación industrial, siempre que no haga explícita renuncia a este hecho. En este caso, deberá autorizar expresamente los proyectos presentados por otros.

12. El Ingeniero Director del presente proyecto, será el responsable de la dirección de la aplicación industrial siempre que la empresa consultora lo estime oportuno. En caso contrario, la persona designada deberá contar con la autorización del mismo, quien delegará en él las responsabilidades que ostente.