Universidad Autónoma de Madrid
Escuela politécnica superior

Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación

Trabajo Fin de Grado

RECOPILACIÓN Y USO DE DATOS MASIVOS EN SISTEMAS DE VERIFICACIÓN DE FIRMA MANUSCRITA DINÁMICA

Autor: Rubén Barco Terrones
Tutor: Rubén Tolosana Moranchel
Ponente: Julián Fiérrez Aguilar

JUNIO 2018

RECOPILACIÓN Y USO DE DATOS MASIVOS EN SISTEMAS DE VERIFICACIÓN DE FIRMA MANUSCRITA DINÁMICA

Autor: Rubén Barco Terrones
Tutor: Rubén Tolosana Moranchel
Ponente: Julián Fiérrez Aguilar

Biometrics and Data Pattern Analytics - BiDA Lab
Dpto. de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Universidad Autónoma de Madrid
JUNIO 2018

Resumen

Resumen

En este Trabajo Fin de Grado se genera una base de datos de firmas manuscritas on-line de una magnitud mayor a las existentes en el estado del arte y que recoge los escenarios de captura de firma on-line más utilizados en la actualidad, tales como el uso de diferentes dispositivos de captura y útiles de escritura. Dicha base de datos se ha formado a partir de la unificación de algunas de las bases de datos ás utilizadas en el ámbito de investigación. Ha sido necesario para ello realizar una etapa de preprocesado y organización de la memoria utilizando una nueva nomenclatura que permita su fácil uso en trabajos futuros. A su vez, se han estudiado, implementado y evaluado sistemas de reconocimiento biométrico de firma dinámica basados en Redes Neuronales Recurrentes (RRN).

Como punto de partida, se ha realizado un estudio inicial del estado del arte en el ámbito de verificación de firma manuscrita on-line. Al mismo tiempo, se ha profundizado en las técnicas basadas en RNN y en su aplicación en firma manuscrita con arquitecturas siamesas y bidireccionales.

Una vez entendido el estado del arte desde un punto de vista teórico, el siguiente paso ha consistido en definir y programar el diseño de la base de datos. Esta base de datos cuenta con un número de usuarios elevado con el que se pueden obtener unos resultados fiables. En ella se recogen varios escenarios como la multi-sesión, variabilidad del útil de escritura, diferentes dispositivos de captura, etc. Cada usuario tiene un número determinado de firmas genuinas y falsificaciones.

La parte experimental se ha llevado a cabo en tres etapas. En primer lugar se ha evaluado el sistema propuesto basado en RNN utilizando la nueva base de datos BiDA MDI-Sign para el escenario de captura con el stylus, mejorando los resultados obtenidos en trabajos anteriores. En la segunda etapa se ha realizado un análisis de los resultados obtenidos por nuestro sistema propuesto en función de la complejidad de la firma. Por último, se ha estudiado el escenario de firma realizada con el dedo, siendo un estudio novedoso y de gran relevancia en los escenarios actuales.

Finalmente, se presentan las conclusiones extraídas a lo largo de este trabajo, así como los posibles enfoques de trabajo futuro.

Palabras Clave

Sistema biométrico, verificación, firma on-line, BiDA MDI-Sign, arquitectura siamesa, DTW, RNN, BLSTM, BGRU.

Abstract

In this Final Degree Project, a database of online handwritten signatures of a magnitude greater than those existing in the state of the art is generated, which includes the current most widely used on-line capture scenarios, such as the use of different capture devices and writing tools. This database has been formed from the unification of some of the databases used in the field of research. It has been necessary to carry out a stage of preprocessing and organization of the memory using a new nomenclature that allows its easy use in future work. At the same time, dynamic signature biometric recognition systems based on Recurrent Neural Networks (RRN) have been studied, implemented and evaluated.

As a starting point, an initial study of the state of the art has been carried out in the area of online handwritten signature verification. At the same time, the techniques based on RNN and its application in handwritten signatures with bidirectional and siamese architectures have been deepened.

Once understood the state of the art from a theoretical point of view, the next step has been to define and program the design of the database. This database has a high number of users with which you can obtain reliable results. It contains several scenarios such as multi-session, variability of the writing utility, different capture devices, etc. Each user has a certain number of genuine signatures and forgeries.

The experimental part has been carried out in three stages. First of all, the proposed system based on RNN was evaluated using the new BiDA MDI-Sign database for the capture scenario with the stylus, improving the results obtained in previous works. In the second stage an analysis of the results obtained by our proposed system based on the complexity of the firm was made. Finally, the signing scenario with the finger has been studied, being a novel study and of great relevance in the current scenarios.

Finally, the conclusions drawn from this work are presented, as well as the possible future work approaches.

Key words

Biometric system, verification, on-line signature, BiDA MDI-Sign, siamese architecture, DTW, RNN, BLSTM, BGRU.

Agradecimientos

En primer lugar, quiero dar mi más sincero agradecimiento a mi tutor Rubén Tolosana por haberme brindado la oportunidad de realizar este trabajo. Por su atención, apoyo y rapidez a la hora de resolver dudas, por su constante disposición a ayudarme y facilitarme contenido y por sus continuas ganas de enseñar y aprender. También, agradecer a Rubén Vera por la ayuda y las ideas proporcionadas a pesar de no ser mi tutor, y a todo el grupo de investigación BiDA Lab por ponerse en contacto conmigo hace más de un año y permitirme trabajar a su lado.

Me gustaría dar las gracias también a todo mi grupo de amigos de toda la vida. Por entenderme en cada momento que he pasado a su lado y por preocuparse por mí y animarme cuando las cosas no iban del todo bien. Por el simple hecho de quedar a tomar una *cerve*, muchísimas gracias.

Me gustaría dedicar unas líneas a mis compañeros de la carrera. Han sido cuatro años de constante trabajo, exámenes y prácticas. Gracias por haberme enseñado tanto, por haberme aguantado y por haberme permitido conocer a personas tan afines a mí.

Cómo no, agradecer de todo corazón a mi padre, mi madre y mi hermana Beatriz. Sois un apoyo constante. Gracias por aguantar mis malas contestaciones en épocas de exámenes y por el esfuerzo que vais a hacer por mí.

Y por último, dar las gracias a mi compañero Pablo Lázaro. Sin duda ha sido la persona más cercana a mí en estos cuatro años de carrera. Ha sido un gran apoyo para mí tanto en los buenos como en los malos momentos. Por la ayuda que me ha proporcionado en toda la carrera y en este último trabajo. Vales mucho como persona, como amigo y como ingeniero. Sin ti esta etapa hubiera sido muy diferente.

Muchas gracias a todos.

Rubén Barco Terrones

Junio 2018

Índice general

Índice de Figuras

Índice de Tablas

1 Introducción

1.1. Motivación

El gran avance tecnológico producido en los últimos años ha originado como resultado el uso masivo de los dispositivos móviles en la sociedad para todo tipo de tareas, desde realizar pagos on-line a gestiones administrativas. El sector público y privado es consciente de estos importantes cambios en la sociedad y está poniendo todos sus esfuerzos en el desarrollo de nuevos métodos de autenticación que sustituyan las carencias de los métodos tradicionales basados en lo que los usuarios tienen (tarjetas de acceso, tokens, etc.) o conocen (PINs o passwords). Los sistemas de autenticación basados en la información biométrica de la persona permiten solventar dichas carencias al mismo tiempo que proporcionan entornos de aplicación amigables.

La autenticación biométrica es el proceso de verificar la identidad de un individuo a través de sus características físicas o conductuales. En los últimos años su uso se ha extendido a todo tipo de sistemas de seguridad, muchos de ellos ya implantados en la sociedad, como la huella dactilar en los teléfonos móviles, el reconocimiento facial para desbloquear el PC o para controles fronterizos, sistemas de detección de sospechosos por voz o por patrones de comportamiento, etc. Por otro lado, la proliferación en los últimos años de los dispositivos táctiles ha permitido una gran expansión de los sistemas de reconocimiento biométrico basados en la escritura y la firma manuscrita.

La firma manuscrita siempre ha sido el método tradicionalmente utilizado para la autenticación de documentos legales y, a día de hoy, lo sigue siendo. La gran aceptación por parte de la sociedad como método de autenticación hace necesario que se realice un profundo estudio de los nuevos escenarios de implantación, en especial, los escenarios móviles. Pero el reconocimiento biométrico de firma es un problema complejo. Actualmente, un usuario puede realizar su firma en múltiples dispositivos de captura (e.g. smartphones, tablets, etc.) así como con diferentes útiles de escritura (e.g. stylus y dedo). Por otro lado, existe una alta variabilidad entre firmas de un mismo usuario (intra-clase) y una baja variabilidad entre firmas de diferentes usuarios (inter-clase).

Aquí es donde surge la principal motivación de este trabajo: explorar el uso de nuevas arquitecturas y algoritmos basados en redes neuronales profundas, en concreto Redes Neuronales Recurrentes (RNN), en el problema de verificación de firma manuscrita on-line con los dos siguientes objetivos:

1. Mejorar el rendimiento de los sistemas tradicionales de verificación de firma manuscrita on-line.

2. Obtener un sistema que sea capaz de generalizar, es decir, que sea robusto frente a los nuevos escenarios descritos, tales como interoperabilidad de dispositivos, captura de firmas con dedo o stylus, usuarios con distinto grado de complejidad de la firma, etc.

Debido al funcionamiento de los sistemas basados en redes neuronales profundas, que funcionan bien cuando tenemos muchos datos, para poder desarrollar sistemas robustos que extraigan características relevantes independientemente del escenario de captura, se hace necesaria la adquisición de una nueva base de datos extensa que englobe distintos dispositivos y usuarios, lo que supondrá la adaptación de los algoritmos a esta nueva base de datos.

1.2. Objetivos

Este trabajo se ha llevado a cabo con el propósito de cumplir los dos siguientes objetivos:

1. Creación de una nueva base de datos de firma on-line a partir de la recopilación de las bases de datos más importantes en el ámbito. Esto permitirá el desarrollo e investigación de nuevos sistemas en los nuevos escenarios descritos, tales como interoperabilidad de dispositivos y útiles de escritura, escenarios multi-sesión, variabilidades inter-clase e intra-clase, usuarios de distinta complejidad y diferentes tasas de muestreo.

2. Explorar el uso de sistemas basados en redes neuronales profundas, concretamente RNN, en el ámbito de la firma manuscrita on-line con el objetivo de ver su potencial en comparación con los sistemas tradicionalmente utilizados. Estos algoritmos están siendo ampliamente utilizados en otro tipo de tareas como reconocimiento de voz o escritura con resultados excelentes. Sin embargo, la escasez de datos, así como el escenario de verificación y no de identificación, ha originado por el momento escasos resultados en el ámbito de verificación de firma manuscrita on-line. Por este motivo, el segundo de los objetivos es desarrollar un sistema de verificación de firma on-line basado en RNN que sea capaz de generalizar frente a los nuevos escenarios descritos.

Los pasos necesarios que se han propuesto para cumplir con los objetivos son los siguientes:

- Creación de una base de datos de firma on-line de mayor magnitud que las actuales y en la que se tengan en cuenta los diferentes escenarios comentados, incluyendo dispositivos específicos para la captura de firma y escritura (tablets Wacom) y dispositivos de uso genérico (tablets y smartphone Samsung). Cabe destacar la importante labor de organizar esta gran base de datos de manera que precise de una fácil accesibilidad para estudios futuros y de una nomenclatura clara y estricta.

- Comprensión del código que permita la realización de experimentos con esta base de datos y estudio del estado del arte de los sistemas de verificación de firma on-line.

- Adecuación de las arquitecturas utilizadas en trabajos previos, así como exploración de los parámetros óptimos para los escenarios descritos anteriormente (e.g. multi-sesión, multi-dispositivo, etc.).

1.3. Metodología y plan de trabajo

Para alcanzar los objetivos establecidos en este Trabajo Fin de Grado, se ha seguido la metodología que se muestra en la Fig. 1.1, detallada a continuación.

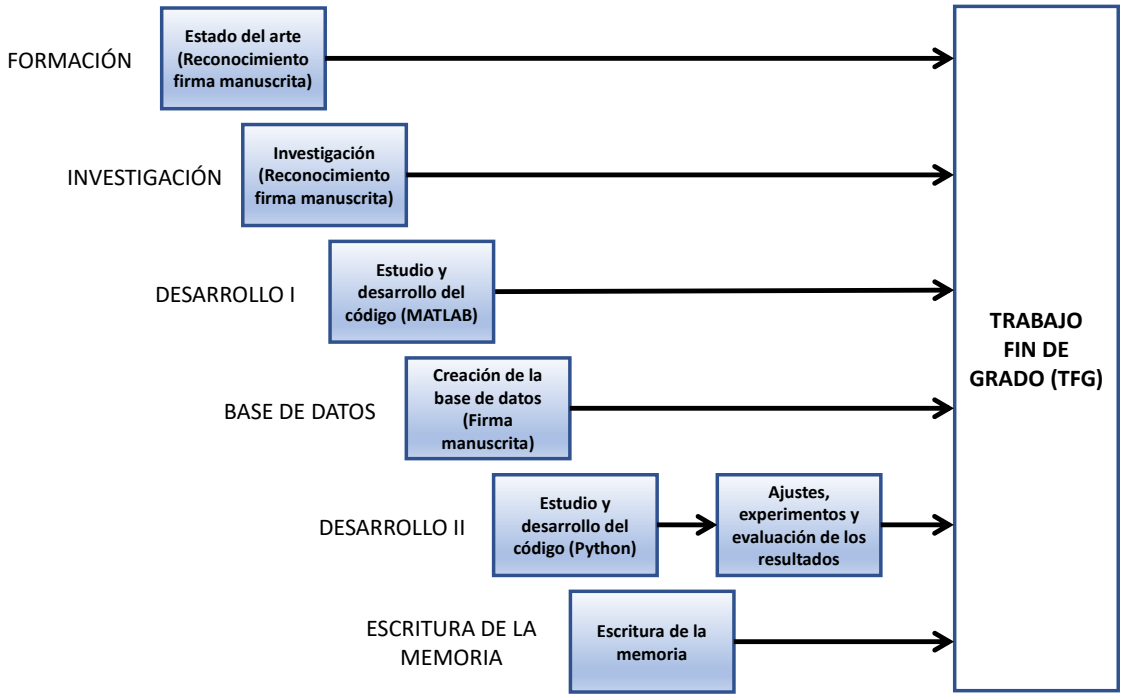


Figura 1.1: Diagrama del plan de trabajo seguido.

- **Estudio del estado del arte.** A la hora de llevar a cabo un proyecto, lo primero que se debe hacer es obtener la información y los conocimientos necesarios para realizarlo. En este trabajo en concreto, en primer lugar, se ha llevado a cabo un estudio de los conceptos más importantes del reconocimiento biométrico. Además, se ha estudiado el estado del arte de los sistemas de verificación de firma on-line, utilizando libros, publicaciones y cursos on-line.

- **Estudio y desarrollo del software (I).** El primer paso para conseguir reunir varias bases de datos en una sola es realizar una etapa de preprocesdo. Además, se debe otorgar, tanto a usuarios como a sus correspondientes firmas, una nomenclatura que sea común para toda la base de datos final. Para poder realizar todo este proceso, se ha llevado a cabo una etapa de estudio y programación del código a implementar.

- **Creación de la base de datos.** La primera parte de este trabajo se ha basado en la realización de una base de datos que engloba todos los escenarios actuales de adquisición de firma manuscrita tales como multi-dispositivo, multi-sesión y captura de firmas con diferentes útiles de escritura entre otros, como se ha descrito en los objetivos. Se ha estudiado el modo de organizar esta base de datos para que proporcione una fácil accesibilidad y un cómodo uso. Esta primera parte del proyecto se ha llevado a cabo mediante la herramienta **Matlab**.

- **Estudio y desarrollo del software (II).** Después de tener organizada la base de datos, se realizado un proceso de familiarización con las herramientas y programas disponibles

en el grupo BiDA Lab para llevar a cabo la experimentación posterior. En este caso, se trata de la comprensión del software utilizado (e.g. Matlab, Keras, Theano, TensorFlow) para la comprensión del código ya existente en el grupo, así como la el desarrollo de nuevo código que permita alcanzar los objetivos marcados en este Trabajo Fin de Grado.

- **Desarrollo y evaluación de los experimentos.** Posteriormente, se han realizado experimentos con diferentes técnicas y siguiendo los protocolos de evaluación utilizados en trabajos anteriores con el objetivo de mejorar los resultados del estado del arte. Además, se han analizado y evaluado los resultados obtenidos para ver el efecto que tienen los diferentes escenarios explicados anteriormente en los sistemas entrenados, todo ello en términos de *Equal Error Rate* (EER). El código utilizado está organizado y comentado para un posible uso posterior.
- **Escritura de la memoria.** Finalmente, tras llevar a cabo los diferentes experimentos y analizar y evaluar los resultados obtenidos, se procede a realizar un estudio y una comparación con los resultados del estado del arte para desarrollar la memoria del presente Trabajo Fin de Grado.

1.4. Organización de la memoria

El presente Trabajo Fin de Grado consta de seis capítulos:

- **Capítulo 1: Introducción**
- **Capítulo 2: Estado del arte**
- **Capítulo 3: Base de datos**
- **Capítulo 4: Sistema propuesto**
- **Capítulo 5: Desarrollo experimental**
- **Capítulo 6: Conclusiones y trabajo futuro**

2

Estado del arte

2.1. Importancia de la biometría en la actualidad

La sociedad actual se encuentra en un punto en el que existe una gran globalización e interconectividad. Este es el motivo por el cual la tecnología de los sistemas de verificación de usuarios está evolucionando y expandiéndose rápidamente. La biometría se usa actualmente en muchos ámbitos de la sociedad, entre los que destacan los siguientes:

- **Telefonía**: A lo largo de los últimos cinco años, ha predominado la implementación de lectores de huellas dactilares en los teléfonos móviles. Además, existen compañías que han introducido el desbloqueo de dispositivos móviles por reconocimiento facial, sustituyendo así métodos tradicionales y más vulnerables como el patrón de desbloqueo o la contraseña. Recientemente se han llevado a cabo estudios sobre autenticación de personas mediante la interacción del usuario con la pantalla [1] [2].
- **Banca**: Gracias a la implementación de los sistemas de huella digital en los teléfonos móviles, muchos compañías bancarias permiten confirmar pagos o transacciones desde el propio dispositivo a través de la huella dactilar. Este sector está realizando una importante inversión en sistemas de reconocimiento biométrico debido a la importancia de la seguridad para evitar grandes pérdidas económicas. Actualmente, está creciendo el interés del sector bancario en los sistemas de verificación de firma manuscrita debido a que es el método de validación de documentos más extendido en todo el mundo.
- **Ámbito forense**: En los ámbitos de criminalística y procesos jurisdiccionales ha sido muy importante el uso de la huella dactilar para la autenticación de sospechosos. Pero en los últimos años ha aumentado el número de rasgos biométricos utilizados en este sector. Recientemente se ha desarrollado una herramienta diseñada para llevar a cabo el análisis de firmas dinámicas y dar apoyo a las técnicas tradicionales de los peritos caligráficos. [3].

2.2. Características de los sistemas biométricos

Los sistemas de reconocimiento biométrico se basan en identificar patrones para poder llevar a cabo el reconocimiento de personas en campos muy variados. Estos patrones se extraen de

Tipos de rasgos biométricos	
Físicos	**De Comportamiento**
Cara	Escritura
Geometría de la mano	Modo de andar
Iris	Gestos
Venas de retina	Firma manuscrita
Voz	Voz

Cuadro 2.1: Tipos de rasgos biométricos.

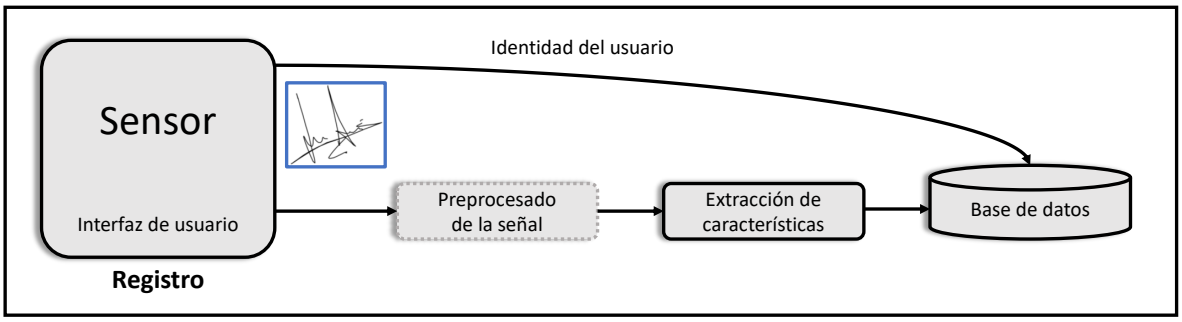


Figura 2.1: Esquema del funcionamiento de un sistema de reconocimiento biométrico. Etapa de registro/Fase de entrenamiento. Figura adaptada de [4].

ciertos rasgos biométricos, que se pueden clasificar en rasgos **físicos** (e.g. la cara) o rasgos **de comportamiento** (e.g. la firma), ver Tabla 2.1. La principal ventaja de la biometría es que, en general, no puede ser olvidada, robada o perdida.

Para que los rasgos mencionados anteriormente puedan ser calificados como biométricos, tienen que cumplir los siguientes requisitos [4]:

- *Universalidad*: todas las personas tienen que tener esta característica.
- *Distintividad*: dos personas cualquiera deben ser suficientemente diferentes en lo que al rasgo biométrico se refiere.
- *Permanencia*: la característica debe ser invariante durante un periodo de tiempo.
- *Mensurabilidad*: el rasgo debe ser cuantificable.

Un sistema biométrico se considera práctico cuando sus rasgos biométricos cumplen también otra serie de características. En primer lugar, deben tener buen **rendimiento**, tanto en las tasas de error como en el proceso de adquisición de los mismos. Otra característica es la **aceptabilidad** del sistema por parte de la sociedad para su uso diario. Por último, los rasgos biométricos deben proporcionar cierta **robustez** a métodos fraudulentos o falsificaciones.

Aun así, no existe ningún rasgo biométrico que cumpla todas las características anteriores. Todos ellos presentan sus ventajas y desventajas, como se puede apreciar en la Tabla 2.2.

Dependiendo del contexto en el que se vaya a realizar la aplicación, existen sistemas biométricos de verificación y sistemas biométricos de identificación. Previamente al funcionamiento del sistema, hay una etapa de **registro**, ver Fig. 2.1.

En esta etapa se recopila la información de los usuarios. Los rasgos biométricos se capturan usando sensores. Una vez capturada esta información se pasa a una etapa opcional de

Rasgo Biométrico	Universalidad	Unicidad	Permanencia	Mensurabilidad	Rendimiento	Aceptabilidad	Evitabilidad
ADN	A	A	A	B	A	B	B
Oreja	M	M	A	M	M	A	M
Cara	A	B	M	A	B	A	A
Termograma facial	A	A	B	A	M	A	B
Venas de la mano	M	M	M	M	M	M	B
Huella dactilar	M	A	A	M	A	M	M
Forma de andar	M	B	B	A	B	A	M
Geometría de la mano	M	M	M	A	M	M	M
Iris	A	A	A	M	A	B	B
Huella palmar	M	A	A	M	A	M	M
Olor	A	A	A	B	B	M	B
Retina	A	A	M	B	A	B	B
Firma	B	B	B	A	B	A	A
Forma de teclear	B	B	B	M	B	M	M
Voz	M	B	B	M	B	A	A
Escritura	B	B	B	A	B	A	A

Cuadro 2.2: Comparación cualitativa de varios rasgos biométricos. A=Alto, M=Medio, B=Bajo.

pre-procesado y después se extraen las características de cada uno de lo usuarios, que se almacenan en una base de datos. Dependiendo del tipo de sistema, estas características se utilizan como plantilla para realizar comparaciones o como datos para entrenar un modelo estocástico de la identidad a reconocer. En el segundo caso, a esta etapa también se le llama **fase de entrenamiento**.

Una vez se ha recopilado la información de los usuarios, tenemos dos tipos de sistemas biométricos:

1. Sistemas de **identificación**: el sistema valida la identidad de un usuario comparando sus rasgos con los datos almacenados en la base de datos, es decir, trata de encontrar la identidad del usuario del cual se ha captado cierto rasgo a través de un sensor. Es una comparación uno-a-muchos, por lo que supone un coste computacional bastante alto, aumentando de forma lineal con el tamaño de la base de datos. Estos sistemas proporcionan la identidad del usuario como salida, en caso de que este se encuentre en la base de datos, o un mensaje indicando que el usuario no pertenece a dicha base de datos, ver Fig. 2.2.

2. Sistemas de **verificación**: los usuarios se identifican con una tarjeta, un PIN o un nombre de usuario. El sistema obtiene el rasgo biométrico del usuario a través de un sensor y compara uno-a-uno con los patrones de ese usuario almacenados en la base de datos. La salida del sistema consiste en la identificación positiva, en el caso de que el sistema detecte cierto nivel de similitud entre los rasgos comparados, o negativa, en el caso contrario, ver Fig. 2.3. La verificación de identidad se suele utilizar para evitar que varias personas usen la misma identidad. Este tipo de sistemas son los que se van a estudiar en este proyecto.

Los sistemas de verificación se basan en dos tipos de errores:

- **Falsa Aceptación (FA)**: error producido cuando un usuario se intenta hacer pasar por otro y el sistema lo reconoce como genuino.
- **Falso Rechazo (FR)**: error producido cuando un usuario genuino es rechazado por el sistema como si fuera un usuario impostor.

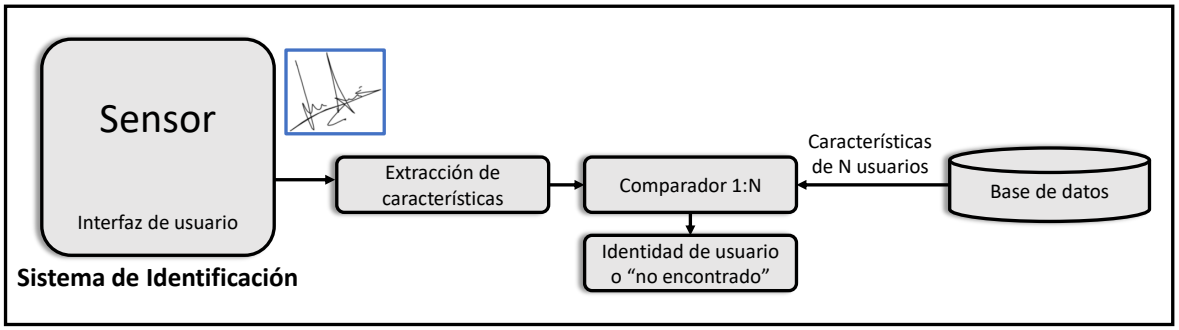


Figura 2.2: Esquema del funcionamiento de un sistema de reconocimiento biométrico. Modo de Identificación. Figura adaptada de [4].

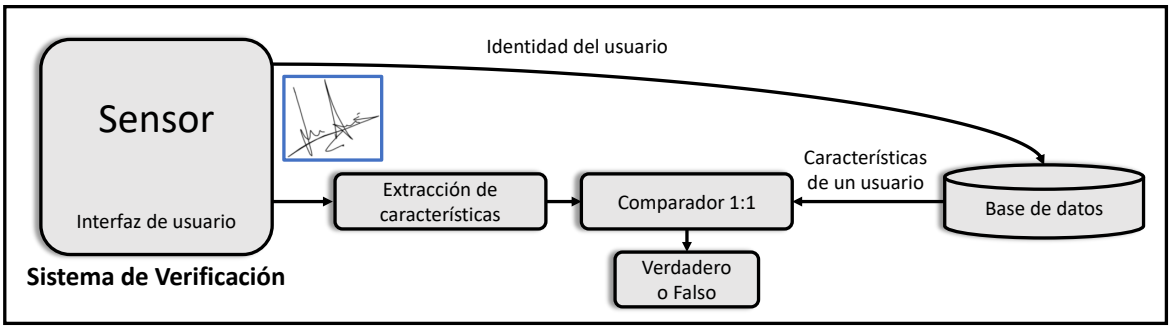


Figura 2.3: Esquema del funcionamiento de un sistema de reconocimiento biométrico. Modo de Verificación. Figura adaptada de [4].

A partir de estos dos errores, del entrenamiento de un sistema de verificación biométrico y de un umbral de decisión, se puede obtener la Tasa de Falsa Aceptación (FAR) y la Tasa de Falso Rechazo (FRR). Con ellas, aparece un nuevo método de medida del rendimiento del sistema, que aporta más información que las tasas de acierto o de error. Se trata del *Equal Error Rate (EER)*, que consiste en la tasa de error del sistema a un determinado umbral de decisión cuando se cumple la condición FAR = FRR. En este trabajo se va a hacer uso de las curvas *DET (Detection Error Trade-off)*, ver Fig. 2.4, que permiten representar adecuadamente el rendimiento de los sistemas de verificación de firma, ya que son sistemas cuya salida se trata de una decisión binaria.

2.3. Redes Neuronales

Las **Redes Neuronales** (*Neural Networks, NN*) son sistemas formados por varios elementos de procesamiento simples, altamente conectados, que procesan la información de entrada a través de su respuesta de estado dinámico [5] [6]. Están inspiradas en la estructura y funcionamiento del córtex cerebral, pero a menor escala y con conexiones más sencillas. Las redes neuronales tienen tres características fundamentales:

1. *Auto-Organización y Adaptabilidad*: utilizan algoritmos de aprendizaje adaptativo y auto-organización permitiendo un procesado más robusto.

2. *Procesado no Lineal*: permite a la red aumentar su inmunidad frente al ruido y aumentar su capacidad de aproximar funciones o de clasificación.

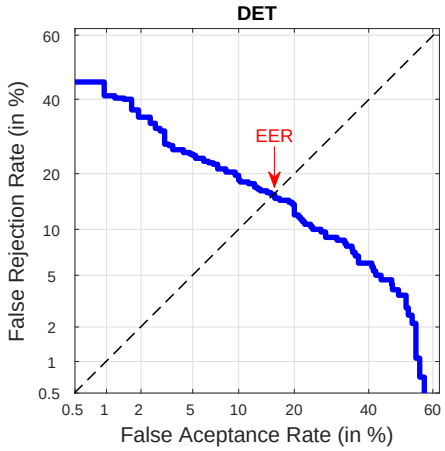


Figura 2.4: Ejemplo de una curva DET.

3. *Procesado Paralelo*: utiliza un gran número de nodos con un alto nivel de interconectividad.

Las neuronas se suelen denominar *nodos* o unidades. Reciben varios inputs (x_j) y producen un valor de salida en función de unos pesos (w_j) asociados a cada input. El valor de estos pesos se va modificando en el proceso de aprendizaje. Cada nodo aplica una función que suma los inputs ponderados mediante los pesos. También se define un valor de umbral o *threshold* a partir del cual se decide el valor de la salida, ver Fig. 2.5(a). Dependiendo de estos parámetros existen diferentes tipos de nodos. Las primeras nodos fueron los denominados perceptrones [5]. Se trata de nodos muy sencillos cuyos valores de entrada y salida sólo podían ser 0 ó 1, limitando mucho su uso si se compara con los nodos utilizados actualmente. Su rendimiento mejoró considerablemente con la introdución de la función sigmoidal de activación, con la que estos valores de entrada y de salida podían ser cualquier valor comprendido en el intervalo de 0 a 1 y se calculaban de la siguiente manera:

$$\sigma(w \cdot x + b) = \frac{1}{1 + exp(-\sum_j w_j x_j - b)} \tag{2.1}$$

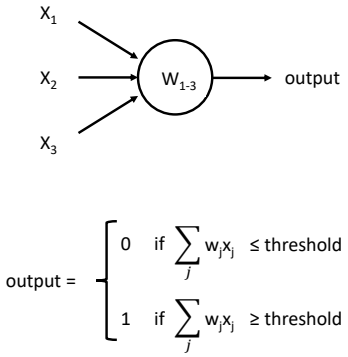
donde b corresponde a un valor de *bias*. Otras funciones de activación muy importantes en la actualidad son la función Gaussiana, la tangente hiperbólica y la función ReLU (*Rectified Linear Unit*).

Las características de los nodos son propicias para el procesado de imágenes y de señales temporales. La combinación entre ellos genera una arquitectura que combina procesado adaptativo paralelo con interconexiones jerárquicas. Consta de dos fases:

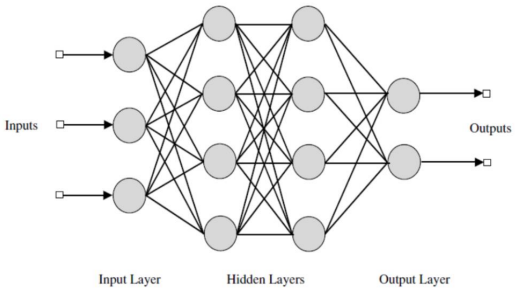
- **Fase de entrenamiento**: se utiliza un conjunto de datos para que la red vaya adaptando sus pesos lo mejor posible definiendo un modelo de red neuronal. Los pesos se van calculando de manera iterativa, con el objetivo de minimizar una **función de coste** que mide el error entre la salida real que se desea obtener y la salida de la red neuronal.

- **Fase de prueba**: es posible que el modelo fruto de la fase de entrenamiento se haya adaptado demasiado a las características de ese conjunto (i.e. *overfitting*), no siendo capaz de generalizar su aprendizaje a casos nuevos. Para evitar este sobreajuste, se utiliza otro conjunto de datos diferentes a los de entrenamiento, para controlar el proceso de aprendizaje de la red. Este conjunto de datos se llama conjunto de *validación*.

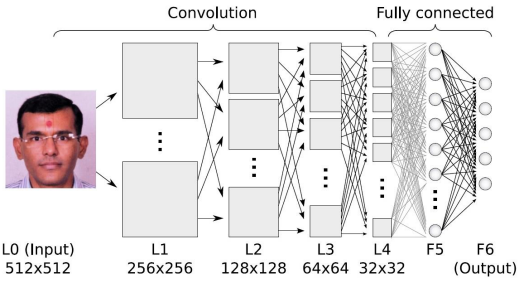
El hecho de utilizar datos etiquetados o no etiquetados durante el entrenamiento da lugar a dos tipos de redes: redes neuronales supervisadas y no supervisadas. Las **redes neuronales de**



(a) Ejemplo de neurona básica de 3 entradas y expresión que computa el valor de salida.



(b) *Multi Layer Perceptron* (*MPL*).



(c) Arquitectura de Red Neuronal Convolucional para reconocimiento facial.

Figura 2.5: Redes Neuronales Artificiales.

aprendizaje supervisado son las más comunes. De los datos utilizados para el entrenamiento se conoce tanto la entrada como la salida, por lo que la red puede beneficiarse de conocer el valor que debería dar a su salida. En cambio, en las **redes neuronales de aprendizaje no supervisado** sólo se conocen los patrones de entrada, por lo que la red tiene que adaptarse mediante las experiencias de entrenamientos anteriores. Un método para entrenar este último tipo de redes es la *Regla del Aprendizaje Competitivo*, es decir, si un dato o patrón pertenece a una clase que ya es conocida, los pesos de la red no variarán considerablemente, pero si el nuevo dato no pertenece a ninguna clase conocida, la red deberá ajustarse para reconocer esta nueva clase.

La interconexión entre los nodos da lugar a un conjunto de capas dentro de la arquitectura de la red neuronal, que también supone la aparición de diferentes tipos de redes neuronales. Las capas y los nodos están conectadas por una red de pesos de conexión, que puede ser de 4 tipos (i.e. hacia delante, hacia atrás, lateral o de retardo. Esto da lugar a dos tipos de redes neuronales:

- **Redes monocapa**: los nodos se conectan lateralmente. Algunos de ellos permiten conexiones consigo mismos, dando lugar a redes monocapa autorecurrentes. Las más representativas son la red de Hopfield, la red de memoria asociativa y las máquinas estocásticas de Boltzmann [7] y Cauchy.

- **Redes multicapa**: los nodos están unidos por diferentes tipos de conexiones. Como se puede observar en la Fig. 2.5(b), las redes neuronales multicapa suelen tener tres capas: la capa de entrada (*input layer*), la capa oculta (*hidden layer*) y la capa de salida (*output layer*). Lo más común es que las conexiones entre las capas sean hacia delante (conexiones

feedforward), pero también existen redes en las que puede haber conexiones de capas hacia atrás (conexiones *feedback* o retroalimentadas).

Pero estas no son las únicas clasificaciones que se pueden hacer. A continuación, se van a definir las redes neuronales más utilizadas en la actualidad:

- **Redes Neuronales Convolucionales (CNN)**: son un tipo especializado de redes neuronales que han conseguido un gran éxito en diferentes ámbitos. Su nombre indica que utilizan una operación matemática llamada **convolución** en al menos una de sus capas [5], ver Fig. 2.5(c). Los datos que utiliza pueden ser series temporales, pensados como datos de una dimensión tomados regularmente cada cierto periodo de tiempo, o imágenes, pensadas como una rejilla de dos dimensiones de píxeles. Una de las características más importantes de las CNN es que en cada capa se utiliza una función denominada *Pooling*. Esta función lo que hace es reemplazar la salida de la red en una cierta posición por el resultado de aplicar una operación a los valores vecinos. Por ejemplo, una operación de este tipo puede ser devolver la media ponderada de los vecinos próximos a esa posición. Las Redes Neuronales Convolucionales son muy importantes en la biometría actual, consiguiendo muy buenos resultados en ámbitos muy diferentes como reconocimiento facial [8], reconocimiento espacial de pasos [9] o reconocimiento de escritura [10].
- **Redes Neuronales Recurrentes (RNN)**: son un tipo específico de arquitectura Deep Learning (DL) que está siendo cada vez más importante en la actualidad para modelar secuencias temporales de datos. Su rango de aplicación es muy variado, abarcando campos desde el reconocimiento de voz [11] a problemas biomédicos [12]. Son modelos con capas ocultas (*hidden layers*) autoconectadas. Un beneficio de este tipo de redes neuronales es que la memoria de las entradas previas se mantiene en el estado interno de la red, permitiendo hacer uso de la información del pasado. Este tipo de redes se explicará con mayor precisión en la Sección 2.4.3.

2.4. Sistemas basados en firma manuscrita on-line

2.4.1. Introducción

Como se ha podido observar en la Tabla 2.2, la firma manuscrita es un rasgo biométrico que tiene unas ventajas aprovechables para desarrollar un sistema de verificación. En primer lugar, tiene un alto nivel de aceptabilidad debido a que la firma ha sido comúnmente utilizada como método de validación y autenticación de documentos financieros y legales. Por otro lado, tiene un alto nivel de mensurabilidad, ya que hoy en día existen un gran número de documentos firmados de los que se pueden escanear firmas, así como multitud de dispositivos móviles (e.g. smartphones, tablets) en los que se pueden firmar documentos. Pero el reconocimiento de usuarios mediante sistemas de verificación de firma manuscrita sigue siendo un reto en la actualidad. Esto se debe a tres puntos clave:

- **Alta variabilidad intra-clase**: la firma es un rasgo que está influenciado por las condiciones físicas y emocionales en las que se encuentra el firmante. Esto introduce variabilidades intra-clase entre firmas genuinas de un mismo usuario, que debe tenerse en cuenta a la hora de verificar su identidad.
- **Baja variabilidad inter-clase**: los sistemas de verificación de firma deben tener en cuenta posibles intentos de falsificación, que dan lugar a firmas muy parecidas a las genuinas.

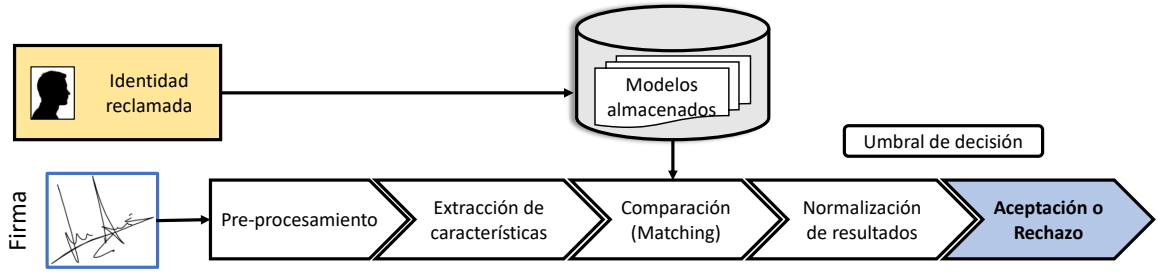


Figura 2.6: Arquitectura típica de un sistema de verificación de firma.

- **Baja permanencia**: la firma suele ser rasgos biométricos que varía con el tiempo. Esto también añade variabilidades intra-clase entre firmas genuinas de un mismo usuario.

A su vez, existen dos tipos de sistemas de verificación de firma en función de qué información se extrae y se utiliza de la firma:

- **Sistemas de verificación de firma off-line**: se extraen las características a partir de las imágenes de las firmas estáticas.
- **Sistemas de verificación de firma on-line**: es el sistema estudiado en este Trabajo Fin de Grado. En este tipo de sistemas, las firmas son capturadas con distintos dispositivos electrónicos (e.g. smartphone, tablets), permitiendo la obtención de la información biométrica del usuario durante todo el proceso de realización, como pueden ser las coordenadas X e Y, la presión ejercida sobre el dispositivo, la inclinación del bolígrafo, etc. Trabajos recientes han probado que estos sistemas consiguen mejores resultados que los sistemas off-line, ya que se tiene una mayor cantidad de información de cada firma.

Otro rasgo importante de la firma manuscrita es la complejidad. Existen firmas de baja, media y alta complejidad. En estudios previos se ha demostrado que los sistemas de verificación de firma son muy sensibles a la complejidad de la firma [13]. Pero este no es el único escenario de estudio posible. Escenarios actuales como interoperabilidad de dispositivos o uso del dedo como útil de escritura suponen un gran desafío para los sistemas de firma manuscrita on-line.

2.4.2. Sistemas tradicionales

Existe una arquitectura básica para los sistemas de verificación de firma manuscrita dinámica. Esta arquitectura puede ser modificada en función del estudio que se vaya a realizar, pero sus elementos típicos son los que se muestran en la Fig. 2.6, explicados a continuación:

1. **Captura de datos**: dependiendo del dispositivo de captura se va a tener mayor o menor cantidad de información almacenada. Los dispositivos de uso genérico (e.g. tablet Samsung con stylus) permiten capturar las coordenadas y la presión de la firma. En cambio, los dispositivos de captura específicos para la adquisición de escritura o firma manuscrita on-line (e.g. tablets Wacom) pueden capturar también el ángulo de inclinación del bolígrafo durante el periodo de firma y la trayectoria del bolígrafo durante los pen-ups (periodo de tiempo en el que se levanta el útil de escritura de la pantalla entre trazos de la firma). Además, la frecuencia de muestreo en las tablets Wacom es Hardware, mientras que la del resto de dispositivos es Software y no tiene tantan calidad y precisión. La información de los pen-ups resulta de gran importancia, como se muestra en varios estudios del estado del

arte [14] [15]. Las señales se muestrean temporalmente a frecuencias entre 100 y 200Hz, cumpliendo con el criterio de Nyquist, ya que la mayor frecuencia observada a la hora de realizar una firma es de 20-30Hz [16].

2. **Etapa de preprocesamiento**: una vez capturados los datos, suele haber un preprocesamiento para filtrar ruido, diezmar la señal discreta para eliminar muestras repetidas, eliminar ceros iniciales y finales, etc. La presencia finales se debe al tiempo que transcurre en el proceso de captura desde que el firmante termina hasta que el operario acepta la firma. Los ceros iniciales suceden por lo mismo pero cuando el firmante comienza a firmar. En este Trabajo Fin de Grado se ha dado mucha importancia a esta parte debido a la búsqueda de coherencia a la hora de crear la gran base de datos que se explica en la Sección 3. Con este preprocesado, se pueden conseguir mejoras notables en la problemática de interoperabilidad de dispositivos.

3. **Extracción de características**: es el proceso de obtención de características discriminatorias. Deriva en dos sistemas: 1) Sistemas basados en características globales (e.g. número de pen-ups, duración o velocidad de la firma) y 2) Sistemas basados en funciones temporales (e.g. presión, trayectoria). El rendimiento de un sistema se degrada cuando los datos de entrenamiento son pequeños comparados con el número de dimensiones del vector de características. Esto suele ocurrir en los sistemas de verificación de firma. Los algoritmos de selección de características se centran en reducir el tamaño del vector de características para optimizar el rendimiento del sistema para un criterio dado (como por ejemplo EER). El caso ideal sería comprobar todos los vectores de características posibles, pero no es realizable computacionalmente, por lo que existen métodos más eficaces como ***Filter method*** y ***Wrapper method***. Existen muchos tipos de algoritmos de selección de características (i.e. ***Scalar Feature Selection***, ***Sequential Forward/Backward Selection***), pero todos llegan a soluciones subóptimas para el problema.

4. **Registro**: existen dos tipos de registros. En el registro basado en modelo se obtiene un modelo estadístico de cada usuario a partir de un conjunto de firmas genuinas de entrenamiento. En el registro basado en referencia se almacenan las características de cada una de las firmas como plantillas y después se compara la firma entrante con el conjunto de plantillas guardadas para ese usuario.

5. **Cálculo de similitud (*matching*)**: los sistemas basados en características globales suelen usar técnicas de medida de distancias (e.g. distancia euclídea o Manhattan). La distancia ecuclídea mide la distancia entre dos puntos en base al Teorema de Pitágoras. La distancia Manhattan, en cambio, nos dice cuál es la distancia entre esos dos puntos sobre una cuadrícula de líneas rectas. Por otro lado, los sistemas basados en funciones temporales suelen usar técnicas que comparan modelos de firmas, como los tres siguientes:

   - ***Hidden Markov Models* (HMM)**: técnica muy utilizada en biometría, destacando en reconocimiento de voz, aunque ha sido utilizada también en sistemas de reconocimiento de firma dinámica [17]. Es un proceso doblemente estocástico en el que el sistema a modelar se considera un proceso de Markov de parámetros desconocidos. Estos parámetros se determinan mediante observaciones, que son modeladas por GMMs ((*Gaussian Mixture Models*).
   - **Dynamic Time Warping (DTW)**: técnica que permite comparar y encontrar la alineación óptima entre dos secuencias temporales de diferente longitud. Permite encontrar regiones correspondientes entre dos secuencias o estimar la similitud entre ellas. Aunque el algoritmo DTW ha sido de utilidad en muchas disciplinas como el reconocimiento de voz, la extracción de datos, la robótica y la medicina, también se ha utilizado en el campo de verificación de firma on-line [18].

- **SVM con *sparse representation***: las *Support Vector Machines* son un método de aprendizaje para resolver problemas de clasificación y regresión. Con *sparse representation* se pueden solucionar corrupciones en la señal (i.e. ruido, datos atípicos y datos perdidos). Esta técnica, combinada con las propiedades de la DCT (*Discret Cosine Transform*), ha sido utilizada en el ámbito de verificación de firma on-line [19].

6. **Normalización de scores**: etapa muy importante. Tras realizar el *matching*, se suelen normalizar los resultados a un rango común. En [20] se estudian algunas de las técnicas de normalización más utilizadas.

2.4.3. Sistemas basados en Redes Neuronales Recurrentes

Uno de los campos en los que las RNN han causado un mayor impacto es en el reconocimiento de escritura debido a la relación que existe entre los inputs actuales y los contextos pasados y futuros. Pero el rango de información contextual al que puede acceder una RNN estándar es limitado. Para solventar esta limitación, surgen dos arquitecturas RNN nuevas: **Long Short-Term Memory** (**LSTM**) y **Gated Recurrent Unit** (**GRU**). Estas dos arquitecturas han sido utilizadas con éxito en escenarios de escritura on-line y off-line. En [21], los autores proponen un sistema basado en el uso de **Bidirectional LSTM** (**BLSTM**) para reconocimiento de texto, consiguiendo mejorar los resultados del estado del arte conseguidos con sistemas HMM y demostrando que este enfoque es más robusto a cambios en el tamaño del diccionario empleado. Estos resultados han sido obtenidos también para sistemas de identificación del escritor. A pesar de estos buenos resultados, todavía se han realizado pocos estudios aplicando estas nuevas arquitecturas RNN a sistemas de verificación de firma manuscrita. En algunos estudios se han utilizado el mismo número de usuarios para entrenamiento y evaluación, además de sistemas que debían ser entrenados cada vez que un usuario se registraba en la aplicación. Otros estudios se han centrado en entrenar sólo con firmas genuinas o entrenar diferentes redes para cada usuario, es decir, escenarios muy limitados incapaces de generalizar bien. De todos estos estudios se ha podido concluir que los sistemas de RNN estándar aplicados a verificación de firma no son apropiados para la escasa cantidad de datos disponible. Los tipos de RNN más utilizados en el campo de verificaicón de firma on-line son las redes LSTM y GRU. A continuación, se van a explicar en profundidad estas dos arquitecturas junto con las **RNN Bidireccionales** (**BRRN**).

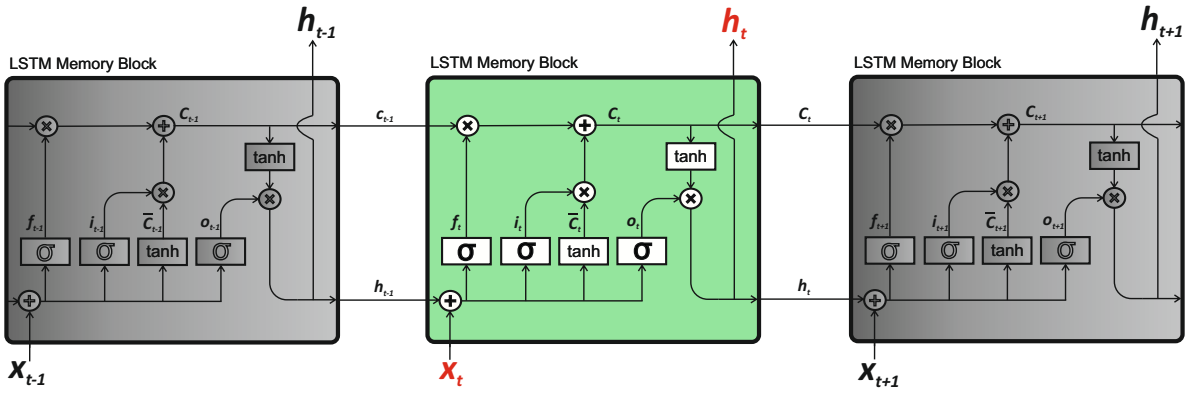


Figura 2.7: Esquema de un bloque de memoria LSTM en diferentes instantes de tiempo (i.e. X_{t-1}, X_t and X_{t+1}). Fuente: [22].

Long Short-Term Memory

Las redes recurrentes LSTM se han utilizado en muchos campos como identificación de lenguaje o problemas biomédicos. Sin embargo, el análisis y diseño de las redes LSTM para nuevos escenarios no es sencillo. Las LSTM RNN están compuestas de bloques de memoria que normalmente contiene una celda de memoria, una *forget gait* f, una *input gate* i y un output gate o. Para cada unidad de tiempo t:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{2.2}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{2.3}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{2.4}$$

$$\widetilde{C_t} = \tanh(W_C x_t + U_C h_{t-1} + b_C) \tag{2.5}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C_t} \tag{2.6}$$

$$h_t = o_t \odot \tanh(C_t) \tag{2.7}$$

donde W_* y U_* son matrices de pesos y b_* es el vector *bias*. El símbolo $\odot$ representa el producto puntual mientras que σ es una capa sigmoidea cuyos valores de salida están entre 0 y 1. Las arquitecturas LSTM tienen la habilidad de borrar información del instante $t-1$ o añadir información nueva en el instante t. La clave es la celda de estado C_t, que está regulada por las diferentes puertas (*gates*). La puerta f decide la cantidad de la información previa que pasa al nuevo estado de la celda C_t. La puerta i indica la cantidad de información a actualizar en la celda de estado C_t. Por último, la salida del bloque de memoria h_t es una versión filtrada de C_t, siendo la puerta o la encargada. En la Fig. 2.7 se muestra un esquema donde se pueden apreciar las relaciones entre las celdas y las puertas.

Las redes LSTM han sido utilizadas recientemente en estudios sobre verificación de firma on-line mejorando los resultados del estado del arte. Esta tipo de RNN se ha utilizado junto con una arquitectura siamesa para mejorar el rendimiento del sistema de verificación [22]. La arquitectura siamesa, que se explicará más adelante en la Sección 4.2, también ha sido implementada en sistemas de verificación de firma off-line junto a CNN [23].

Gated Recurrent Unit

GRU es un nuevo tipo de arquitectura RNN inspirado en LSTM pero mucho más fácil de calcular e implementar. Además, los resultados obtenidos parecen ser muy similares a los obtenidos con sistemas LSTM. La principal diferencia entre ambos reside en la cantidad de puertas

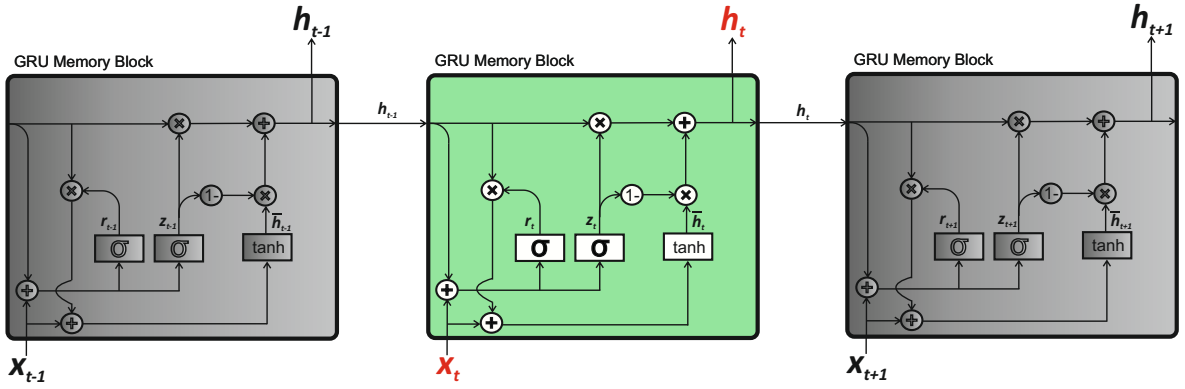


Figura 2.8: Esquema de un bloque de memoria GRU en diferentes instantes de tiempo (i.e. X_{t-1}, X_t and X_{t+1}). Fuente: [22].

utilizadas para controlar el flujo de información. Mientras que las unidades LSTM contienen tres puertas diferentes, las GRU tienen sólo dos puertas: una *reset gate* r y una *update gate* z. Para cada unidad de tiempo t:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{2.8}$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{2.9}$$

$$\widetilde{h_t} = \tanh(W_h x_t + U_h(h_{t-1} \odot r_t) + b_h) \tag{2.10}$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \widetilde{h_t} \tag{2.11}$$

donde W_* y U_* son matrices de pesos y b_* es el vector *bias*. El símbolo $\odot$ representa el producto puntual mientras que σ es una capa sigmoidea cuyos valores de salida están entre 0 y 1. GRU tiene la habilidad de borrar información del instante de tiempo $t-1$ o añadir nueva información nueva en el instante t. La puerta *reset* r_t se encarga de mantener en la celda de estado actual la información del instante de tiempo anterior o reemplazarla por la información de la entrada actual. La puerta de actualización z_t indica cuánta información del instante de tiempo anterior y de la celda de estado actual van hacia la salida del bloque de memoria. En la Fig. 2.8 se muestra un esquema donde se pueden apreciar las relaciones entre las celdas y las puertas.

Lar redes GRU también se han utilizado recientemente en sistemas de verificación de firma on-line obteniendo muy buenos resultados tanto con el uso de descriptores [24] como con arquitecturas siamesas [22].

RNN Bidireccionales

Los esquemas explicados en las dos secciones anteriores son los básicos. Estos esquemas permiten el acceso a la información del pasado y del presente. Pero para algunas aplicaciones como el reconocimiento de escritura o de habla, la oportunidad de tener acceso al contexto futuro puede mejorar el rendimiento del sistema. Los esquemas que también permiten el acceso a la información del futuro se denominan (*Bidirectional RNN*, *BRRN*). Los esquemas BRRN combinan una RNN que se mueve hacia delante en el tiempo empezando por el principio de la secuencia con otra RNN que se mueve hacia atrás en el tiempo empezando por el final de la secuencia, ver Fig. 2.9. Para cada instante de tiempo t, la salida O_t se puede beneficiar de un resumen relevante del pasado y del futuro a la vez.

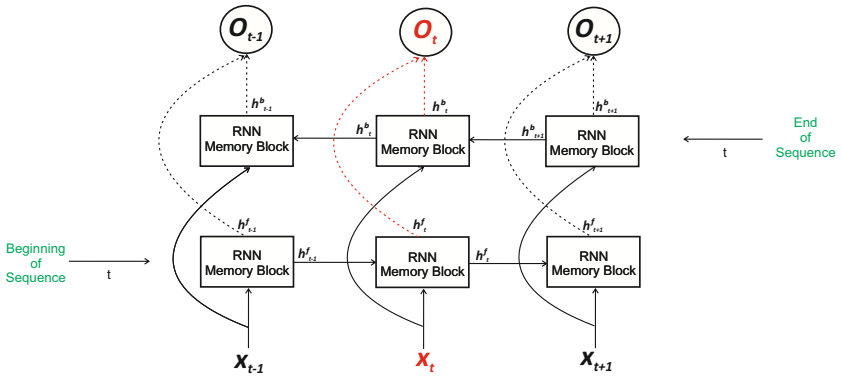


Figura 2.9: Esquema de un sistema RNN bidireccional típico en diferentes instantes de tiempo (i.e. X_{t-1}, X_t and X_{t+1}). Fuente: [22].

3

Bases de Datos

3.1. Introducción

En la actualidad se dispone de varias bases de datos utilizadas en distintos estudios del estado del arte. Esto permite y facilita la comparación de los algoritmos y sistemas implementados. Hace unos años existía una gran carencia de bases de datos públicas debido a los problemas legales, a la privacidad de los usuarios y a las limitaciones tecnológicas para conseguir firmas on-line y off-line, por lo que era complicado crear bases de datos consistentes y que tuvieran en cuenta distintos escenarios. El desarrollo tecnológico ha permitido subsanar estos problemas y conseguir una cantidad de información de firma manuscrita suficiente como para conseguir resultados muy competitivos comparados con otros rasgos biométricos.

Como se ha explicado en la Sección 1.2, el primer paso para alcanzar los objetivos de este proyecto ha sido la creación de una base de firmas manuscritas on-line que contemple los escenarios de adquisición actuales (i.e. interoperabilidad de dispositivos, adquisición de firmas con el dedo y con stylus) y con un volumen de usuarios no visto hasta el momento en ninguna otra base de datos existente. El nombre de la nueva base de datos va a ser **BiDA MDI-Sign (BiDA Lab Multiple Devices and Input Online Signature Database)**.En este punto de la memoria se especificarán las características principales de cada base de datos utilizada, así como el preprocesado que se ha realizado para conseguir la mayor homogeneidad posible. Cabe destacar que se han tenido que realizar modificaciones en algunos usuarios mal capturados o con algún error de captura debido al dispositivo. Por otro lado, se explicarán la nomenclatura y la organización que se han utilizado para unificar todas las bases de datos en una sola.

3.2. BiDA MDI-Sign

3.2.1. Características de las bases de datos utilizadas

Para crear la base de datos extensa, se han utilizado las siguientes bases de datos públicas: MCYT, BiosecurID, Biosecure DS2, e-BioSign DS1, e-BioSign DS2 y e-BioSign DS3. A continuación, se describirán las características principales de cada base de datos: año de captura, número

	MCYT	BiosecurID	BiosecureDS2	e-BioSign DS1	e-BioSign DS2-DS3
Año	2003	2007	2008	2016	2016-2017
Usuarios	330	400	676	65	81
Sesiones	1	4	2	2	2
#muestras genuinas/ usuario/dispositivo	25	16	30	8	8
#falsificaciones/ usuario/dispositivo	25	12	20	6	6
Dispositivo (Útil de escritura)	W. Intous (st)	W. Intuos3 (st)	W. Intuos3 (st)	W. STU-50 (st) W. STU-53 (st) W. DTU-1031 (st) SG. Gal.Note(st/d) SG. ATIV7 (st/d)	W. STU-530 (st) SG. Gal.Note (st/d) SG. Galaxy Neo S3 (d)

Cuadro 3.1: Características de las bases de datos utilizadas. SG=Dispositivo Samsung, W=Tablet Wacom, st=stylus, d=dedo.

de usuarios, número de sesiones, número de firmas genuinas y falsificadas (*skilled forgeries*) y dispositivos y útiles de escritura utilizados, ver Tabla 3.1.

- **MCYT** [25]: la adquisición fue llevada a cabo por diversas instituciones universitarias españolas en el año 2003, entre las que se encuentra el Grupo de Reconocimiento Biométrico ATVS, de la Universidad Autónoma de Madrid. Esta base de datos cuenta con un total de 330 usuarios. Las firmas fueron capturadas usando un dispositivo WACOM Intuos A6 tablet con una frecuencia de muestreo de 100 Hz, permitiendo capturar las coordenadas, la presión y los ángulos del bolígrafo (azimuth y altitud). Hay 25 firmas genuinas y 25 firmas falsificadas por usuario. Las firmas fueron capturadas en grupos de 5. Primero 5 firmas genuinas, luego 5 firmas falsificadas de otro usuario, repitiendo este procedimiento hasta alcanzar las 25 firmas de cada tipo. Cada usuario proporciona 5 firmas falsificadas para los 5 usuarios previos en la base de datos.

- **BiosecurID** [26]: esta base de datos surge como consecuencia del éxito de la base de datos MCYT. Se trata de un proyecto financiado por el Ministerio de Ciencia y Tecnología en el cual han participado seis instituciones académicas españolas entre las que se encuentra el Grupo de Reconocimiento Biométrico ATVS de la UAM. El objetivo es extender la base de datos BIOSEC ya existente, en términos de incluir nuevas sesiones para usuarios ya registrados, así como nuevos usuarios. El número de usuarios final es de 400 y consta de 4 sesiones llevadas a cabo en intervalos de un mes. Una característica de esta base de datos es la distribución balanceada en la edad de los usuarios que participaron, contando con usuarios entre 18 y más de 45 años. Las firmas fueron capturadas con una WACOM Intuos3 A4. En cada una de las 4 sesiones, cada usuario realiza 4 firmas genuinas y 1 firma falsificada para cada uno de los 3 usuarios anteriores. Se consideran 4 escenarios de falsificación.

- **Biosecure DS2** [27]: en este proyecto participan más de 30 instituciones de investigación procedentes de 15 países diferentes. El Grupo de Reconocimiento Biométrico ATVS también participó en la elaboración de dicha base de datos. Fue capturada utilizando una WACOM Intuos3 A6 digitalizada a una frecuencia de muestreo de 100 Hz, con un procedimiento similar al seguido en la base de datos MCYT. El dispositivo captura información de coordenadas de posición, presión y ángulos de inclinación del bolígrafo (azimuth y altitud).

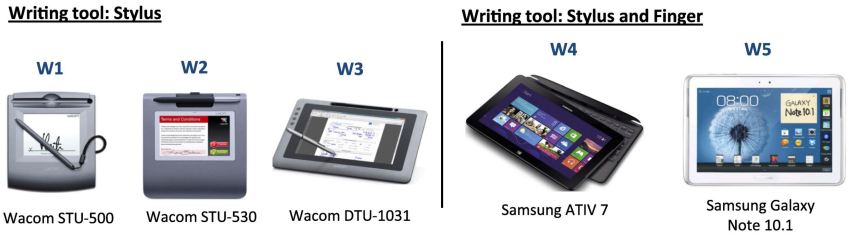


Figura 3.1: Dispositivos de captura utilizados en e-BioSign DS1. Fuente: [18].

Cuenta con un total de 667 usuarios, de 7 países diferentes. Por cada uno de los usuarios se posee un total de 30 firmas genuinas y 20 falsificadas, capturadas en 2 sesiones con un espacio temporal de unos 2 meses. Las firmas fueron capturadas en bloques de 5. En cada sesión los usuarios realizaron 3 sets de 5 firmas genuinas y 5 firmas falsificadas entre cada set. Cada usuario realizó 5 falsificaciones para los 4 usuarios anteriores de la base de datos. El usuario tenía acceso visual a la información dinámica de la firma a falsificar .

- **e-BioSign DS1** [18]: la idea de crear esta base de datos surge al intentar abordar los problemas que supone entrenar y evaluar un sistema con múltiples dispositivos y útiles de escritura. La adquisición fue llevada acabo por el grupo de Reconocimiento Biométrico ATVS exclusivamente. Esta base de datos está compuesta por 65 usuarios, cuyos datos se recopilan en dos sesiones, con un periodo de tiempo entre ellas de 3 semanas. La base de datos está compuesta por cinco dispositivos de captura de escritura a mano. Tres de ellos están diseñados específicamente para esta tarea (dispositivos Wacom), mientras que los otros dos son tabletas de uso general (tabletas Samsung). Vale la pena señalar que los cinco dispositivos se usaron con su propio lápiz óptico y que en los dos dispositivos Samsung se usó también el dedo como herramienta de escritura, ver Tabla 3.1, lo que va a permite analizar el efecto de la herramienta de escritura en el rendimiento del sistema. Además, se utilizó el mismo protocolo de captura para los cinco dispositivos. Los dispositivos Wacom utilizados son los siguientes: Wacom STU-500 (**W1**), Wacom STU-530 (**W2**) y Wacom DTU-1031 (**W3**). Cabe destacar que la frecuencia de muestreo es la misma para todas las Wacom (200Hz) y los niveles máximos de presión varían entre 1024 para STU-530 y 512 para STU-500 y DTU-1031. Los dispositivos genéricos utilizados son: Samsung ATIV 7 (**W4**) y Samsung Galaxy Note 10.1 (**W5**), ver Fig. 3.1. Ambos tienen 1024 niveles de presión. Para cada usuario se capturaron en cada dispositivo 8 firmas genuinas y 6 falsificadas, es decir, 3 genuinas y 2 falsificadas en cada una de las dos sesiones.

- **e-BioSign DS2-DS3**: Esta base de datos fue capturada entre los años 2016 y 2017 por el grupo de Reconocimiento Biométrico ATVS en la Universidad Autónoma de Madrid. En primer lugar se creo solamente la base DS2 con 53 usuarios en 2016 y en 2017 se introdujeron 28 usuarios más, formando un total de 91 usuarios. Los tres dispositivos que han intervenido en esta base de datos han sido una tableta WACOM-STU530 (**W2**), una tableta Samsung Galaxy Note 10.1 (**W5**) y un Smartphone Samsung Galaxy Neo SIII (**W6**). Las firmas fueron realizadas con stylus para W2, con stylus y dedo para W5 y con dedo para W6. En cada dispositivo se realizaron dos sesiones con un intervalo de 15 días entre ellas. El procedimiento para la obtención de genuinas y falsificadas es similar al de e-BioSign DS1.

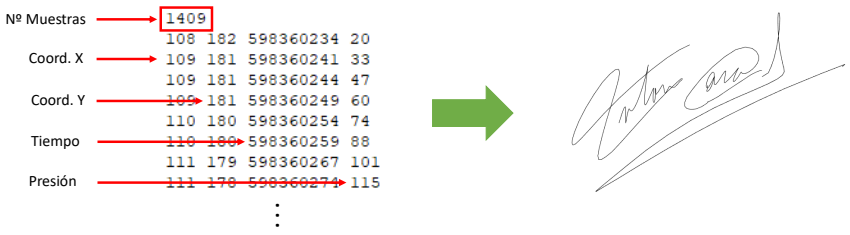


Figura 3.2: Elementos de la firma on-line (izquierda) y su correspondiente firma off-line(derecha).

3.2.2. Preprocesado de las bases de datos

Como se ha explicado en la Sección 2.4.1., la primera etapa de un sistema de verificación de firma manuscrita es el preprocesado de las firmas capturadas. Como el objetivo es crear una base de datos de mayor tamaño que las actuales y que contemple distintos escenarios de adquisición de firma on-line (i.e. interoperabilidad de dispositivos, adquisición de firmas con el dedo y con stylus), se ha de realizar un preprocesado dividido en varios pasos. Todo el preprocesado se ha llevado a cabo sobre la firma on-line, ver Fig. 3.2, es decir, sobre los ficheros con los datos capturados por los dispositivos y programando scripts de Matlab desde cero para cada proceso realizado. Para poder realizar esta etapa, se han organizado las firmas de BiosecureDS2 y BiosecurID por usuarios con el objetivo de partir del mismo esquema que en el resto de bases de datos. Aunque este proceso se realiza de forma independiente en cada una de las bases de datos, existe un conjunto de tareas que se han llevado a cabo en todas por igual y en el siguiente orden:

1. Recorrer todos los usuarios para comprobar que tienen el número de firmas genuinas y falsificadas que corresponde.
2. Búsqueda de firmas vacías en todos los usuarios, es decir, firmas cuyos ficheros no tienen ningún valor almacenado.
3. Conversión de las firmas del formato de origen (i.e. .svc, .fpg) a un formato común para su uso posterior (i.e. .txt).

Una vez realizado este primer preprocesado específico para cada base de datos, se ha estudiado la necesidad de realizar diezmados o de eliminar ceros de presión iniciales y finales en los ficheros de las firmas. El diezmado sólo ha sido necesario en las firmas realizadas con *stylus* sobre los dispositivos W1 y W2 de e-BioSign DS1, debido a que en los estos dispositivos han guardado dos veces seguidas la misma muestra generando ficheros con el doble de longitud. Por este motivo, se ha realizado un diezmado de factor 2. En cuanto a las muestras de presión cero, se producen debido al tiempo transcurrido entre que el operario inicia el proceso de captura y el usuario empieza a firmas, y entre que el usuario termina de firmar y el operario acepta la firma. Este problema se ha tenido que solucionar para varios dispositivos de captura tanto de e-BioSign DS1 como de e-BioSign DS2-DS3 (i.e. W1, W2, W3, W4 y W5).

Por último, se han buscado errores de trazo en las firmas o errores de no correspondencia con el usuario, es decir, errores en la captura de la firma que introducen líneas o puntos no pertenecientes a la misma o firmas que están asociadas a usuarios a los que no corresponden. Para ello se ha realizado una base de datos de firma off-line para cada una de las bases de datos preprocesadas, mediante la representación de las firmas on-line utilizando la información de las coordenadas X e Y (se podría utilizar también la información de presión pero para el objetivo buscado no ha sido necesario) y centrándolas para una mejor visualización. Una vez guardadas

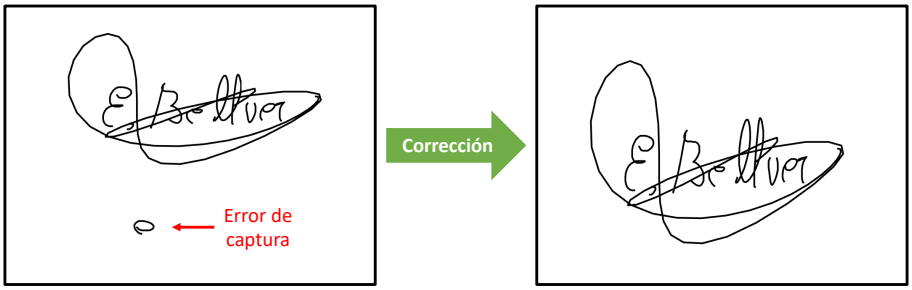


Figura 3.3: Corrección de errores en las firmas.

todas las imágenes se han buscado aquellas en las que existiera algún trazo incorrecto, se ha eliminado esta información de la firma on-line y se ha vuelto a representar para comprobar que el error ha desaparecido, ver Fig. 3.3. Para las firmas que no correspondían con su usuario, se han eliminado o se han sustituido por la firma correcta en el caso de que ésta se encontrara en el directorio de otro usuario.

3.2.3. Organización y nomenclatura

Después de realizar el preprocesado de las firmas y comprobar que todos los usuarios son correctos, se procede a organizar y crear la base de datos definitiva. En este apartado se hablará sobre los usuarios elegidos de cada base de datos individual que formarán parte de entrenamiento o evaluación, teniendo en cuenta que organismo capturó esa base de datos y la cantidad de usuarios disponible en cada una. Como veremos más adelante, la nomenclatura utilizada ha sido diseñada para facilitar el acceso a la nueva base de datos al mismo tiempo que se conserva la información sobre el origen del usuario, dispositivo, sesión de la firma, etc. En cada base de datos se ha seguido la siguiente **organización**, ilustrada en la Tabla 3.2, introduciéndose cada base individual en la base final en el siguiente orden:

- **MCYT**: de los 330 usuarios, 230 se van a utilizar para entrenamiento y 100 para evaluación. La razón de elegir 100 usuarios para evaluación es porque 145 usuarios fueron capturados por el grupo de Reconocimiento Biométrico ATVS, por lo que se podrían hacer públicos en el caso de que alguna entidad los solicite.
- **BiosecurID**: de los 400 usuarios que contiene esta base de datos, 286 han sido seleccionados para entrenamiento y 132 para evaluación, ya que estos 132 usuarios han sido capturados por el grupo de Reconocimiento Biométrico ATVS.

	Development	Evaluation	Total
MCYT	230	100	330
BiosecurID	268	132	400
BiosecureDS2	510	140	650
e-BioSign_DS1	30	35	65
e-BioSign_DS2_DS3	46	35	81
Total	**1084**	**442**	**1526**

Cuadro 3.2: División de los usuarios las bases de datos utilizadas en conjunto de entrenamiento (*development*) y evaluación (*evaluation*).

- **Biosecure DS2**: esta base de datos contiene un total de 676 usuarios, por lo tanto, será la que mas aporte a nivel de número de usuarios y de firmas en la base de datos final. Una vez realizado el preprocesado, como se ha explicado en la sección anterior, se decidió eliminar un total de 26 usuarios incompletos. Finalmente, de los 145 usuarios que capturó ATVS, se eliminaron 5 y se obtuvo un total de 140 usuarios de evaluación. Por otro lado, de los 531 usuarios restantes, se eliminaron 21 usuarios incompletos, quedando un total de 510 usuarios de entrenamiento.

- **e-BioSign DS1**: consta de un total de 65 usuarios. Como esta base de datos ha sido capturada íntegramente por ATVS, la elección de los usuarios ha sido más complicada y se ha intentado que los firmantes pertenecientes a ATVS se encuentren en evaluación. Finalmente se ha decidido que 30 usuarios pertenezcan a entrenamiento y 35 a evaluación.

- **e-BioSign DS2-DS3**: formada por un total de 81 usuarios. Se ha seguido un procedimiento de elección de usuarios similar al seguido en e-BioSign DS1 al tratarse también de una base de datos capturada por ATVS, destinando 46 usuarios a entrenamiento y 35 a evaluación.

Una vez elegido la distribución de cada usuario dentro de la base de datos final, se debe crear una **nomenclatura** para cada sección de la base de datos. Como se muestra en la Tabla 3.2, BiDA MDI-Sign se divide en dos secciones: **Development** (Entrenamiento) y **Evaluation** (Evaluación). Dentro de cada sección, los usuarios van a estar ordenados según la base de datos a la que pertenecen, siguiendo este roden: MCYT, BiosecurID, BiosecureDS2, e-BioSign DS1 y e-BioSign DS2DS3. La nomenclatura utilizada para las carpetas de los usuarios, tanto para evaluation como para development, es la siguiente:

$$u[uid]\_[d/e]\_[bbdd]\_[uid\_bbdd]$$

Donde:

- **uid**: es el número del usuario dentro del conjunto de entrenamiento o evaluación de BiDA MDI-Sign.

- **d/e**: indica si el usuario pertenece al conjunto de entrenamiento (d) o al de evaluación (e).

- **bbdd**: corresponde al nombre de la base de datos de la que procede el usuario.

- **uid_bbdd**: indica el nombre del usuario dentro de la base de datos de la que proviene.

Una vez definido el nombre de la carpeta del usuario, se define una nomenclatura para sus firmas, que es la siguiente:

$$u[uid]\_[g/s]\_[n\_firma]$$

Donde:

- **uid**: es el nombre del usuario dentro del conjunto de entrenamiento o evaluación de BiDA MDI-Sign, al igual que en la nomenclatura diseñada para los usuarios.

- **g/s**: indica si la firma es genuina (g) o una *skilled forgerie* (s) (falsificación).

- **n_firma**: es el nombre original que tenía la firma en su base de datos de origen.

El hecho de utilizar esta nomenclatura va a permitir conocer la información del usuario en su base de datos de origen, pudiendo encontrarlo fácilmente en el caso de pérdida o error. Por otro lado, es sencillo reconocer si una firma es genuina o falsificada, algo muy útil a la hora de realizar los entrenamientos y evaluaciones en los sistemas propuestos más adelante.

Una vez organizada la parte **online** de la base de datos BiDA MDI-Sign con todas las secciones correctamente nombradas y de manera jerárquica, se procede a obtener la misma estructura de secciones para la base de datos **offline**. Para ello, se realiza una copia de la base de datos on-line y se modifica el código empleado en la última parte del preprocesado explicado en la sección anterior para generar las respectivas imágenes. A pesar de que en este Trabajo Fin de Grado no se van a usar las imágenes de las firmas para los procesos de entrenamiento explicados más adelante, es útil disponer de la base de datos off-line para poder realizar un análisis de qué tipo de firmas están fallando en el sistema.

3.2.4. Escenarios considerados

Como se ha explicado en los objetivos de la Sección 1.2, el principal propósito de este Trabajo Fin de Grado es el de conseguir un sistema de verificación de firma manuscrita on-line capaz de generalizar frente a los múltiples escenarios descritos, tales como interoperabilidad de dispositivos, útiles de escritura, usuarios con distinto grado de complejidad de la firma, etc. La base de datos generada tiene en cuenta todos estos escenarios por lo que se pueden considerar los Datasets de la Tabla 3.3. En esta tabla, se han definido dos tipos de dispositivos, A y B. Con A se hace referencia a dispositivos específicos para la captura de firma y escritura (i.e. tablets Wacom) y con B a dispositivos de uso genérico no diseñados para el propósito de adquisición de firma y escritura (i.e. tablets Samsung, smartphones).

Debido a los resultados del estado del arte [22], el primer dataset se ha centrado en el escenario de usuarios de diferente complejidad con firmas realizadas con stylus sobre los dispositivos de captura tradicionalmente urilizados en el ámbito de firma manuscrita (i.e. tablets Wacom), ver Fig. 3.1. En el segundo dataset se ha añadido a la información anterior las firmas capturadas con dispositivos móviles y con stylus, con el objetivo de estudiar la interoperabilidad de dispositivos. En un tercer enfoque, se ha decidido partir de cero con la información de las firmas realizadas con el dedo sobre dispositivos móviles, con el objetivo de comparar los resultados obtenidos al utilizar diferentes útiles de escritura. En este caso sólo se han podido considerar ciertos dispositivos (W4, W5 y W6), presentes en las bases de datos e-BioSign DS1 y e-BioSign DS2-DS3. En los tres primeros Datasets se ha utilizado la información que proporcionan los pen-ups y la presión. Por último, en el Dataset 4 se han considerado todos los escenarios disponibles en la base de

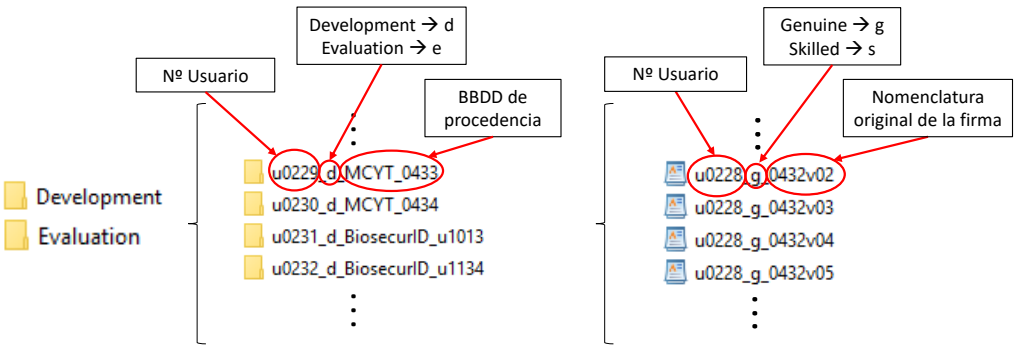


Figura 3.4: Nomenclatura de directorios y firmas de BiDA MDI-Sign.

	DATASET 1	DATASET 2	DATASET 3	DATASET 4
Dispositivos	A	A y B	B	A y B
Útiles de escritura	Stylus	Stylus	Dedo	Dedo y Stylus
Información de pen-ups	Sí	Sí	Sí	No
Información de presión	Sí	Sí	Sí	No
Bases de datos consideradas	MCYT BiosecurID BiosecureDS2 e-BioSign_DS1 e-BioSign_DS2_DS3	MCYT BiosecurID BiosecureDS2 e-BioSign_DS1 e-BioSign_DS2_DS3	e-BioSign_DS1 e-BioSign_DS2_DS3	MCYT BiosecurID BiosecureDS2 e-BioSign_DS1 e-BioSign_DS2_DS3

Cuadro 3.3: Posibles Datasets a considerar.

datos generada, es decir, interoperabilidad de dispositivos y de útiles de escritura y usuarios con diferentes grados de complejidad. Por ello, este Dataset engloba dispositivos Wacom y móviles, así como firmas realizadas con stylus y con el dedo. En este último caso se ha decidido no tener en cuenta la información de la presión y de los pen-ups.

4

Sistema propuesto

4.1. Extracción de funciones temporales

Como se ha descrito en la Fig. 2.6 de la Sección 2.4., uno de los bloques más importantes de un sistema de verificación de firma es la extracción de características. Debido a que se van a hacer uso de RNN, esta extracción de características va a consistir en una serie de funciones temporales. Para cada firma realizada con stylus, los datos correspondientes a las coordenadas X e Y y la presión se han utilizado para extraer un conjunto de 23 funciones temporales, al igual que en [22]. En la Tabla 4.1 se recogen las funciones temporales utilizadas. En cambio, en las firmas realizadas con dedo, sólo se han utilizado los datos de las coordenadas X e Y, ya que la información de presión y su derivada no están disponibles, obteniendo en este caso 21 funciones temporales. Para realizar este procedimiento se ha utilizado la herramienta Matlab.

La información relacionada con el ángulo de orientación del útil de escritura (i.e. ángulos de azimut y altitud) se ha descartado con el objetivo de considerar el mismo conjunto de funciones temporales en todos los dispositivos de captura y útiles de escritura.

#	Feature
1	x-coordinate: x_n
2	y-coordinate: y_n
3	Pen-pressure: z_n
4	Path-tangent angle: θ_n
5	Path velocity magnitude: v_n
6	Log curvature radius: ρ_n
7	Total acceleration magnitude: a_n
8-14	First-order derivate of features 1-7: $\dot{x_n}, \dot{y_n}, \dot{z_n}, \dot{\theta_n}, \dot{v_n}, \dot{\rho_n}, \dot{a_n}$
15-16	Second-order derivate of features 1-2: $\ddot{x_n}, \ddot{y_n}$
17	Ratio of the minimum over the maximum speed over a 5-samples window: v_n^r
18-19	Angle of consecutive samples and first order difference: α_n, $\dot{\alpha_n}$
20	Sine: s_n
21	Cosine: c_n
22	Stroke length to width ratio over a 5-samples window: r_n^5
23	Stroke length to width ratio over a 7-samples window: r_n^7

Cuadro 4.1: Conjunto de funciones temporales consideradas. Fuente: [22].

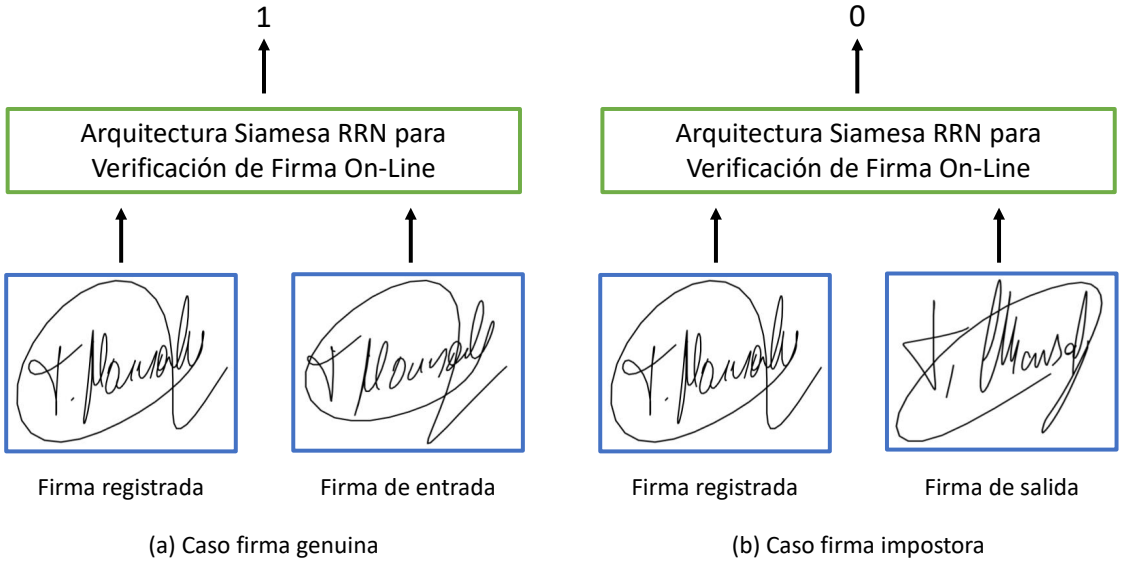


Figura 4.1: Ejemplos de los sistemas RNN propuestos basados en una arquitectura siamesa para minimizar una función de coste. Figura adaptada de [22].

4.2. Sistema basado en RNN

En los sistemas de verificación de firma on-line es común tener un número de firmas de entrenamiento reducido. Una posible solución a este inconveniente es el uso de arquitecturas **siamesas**, más adaptadas al problema de comparación de firmas. Esta arquitectura ha sido utilizada para aplicaciones de reconocimiento o verificación en general. El objetivo principal de esta arquitectura es aprender las similitudes entre muestras de entrenamiento a partir de minimizar una función de coste que conduce a una métrica de similitud grande cuando se trata de pares de firmas genuinas del mismo sujeto y pequeña para pares de firmas de diferentes sujetos. En estudios previos se han propuesto arquitecturas siamesas para CNNs para el problema de verificación facial, obteniendo muy buenos resultados cuando el número de muestras de entrenamiento para una única categoría es pequeño.

En este trabajo se va a implementar un sistema basado en RNN, tomando como base el modelo utilizado en [22], donde se propone por primera vez el uso de las RRN en combinación de una arquitectura siamesa para el escenario de verificación de firma on-line. Con la arquitectura siamesa se consigue que la red aprenda la similitud de pares de firmas. Los autores también proponen sistemas que sean independientes del usuario, es decir, que no tengan que ser reentrenados cuando un nuevo usuario se une a la aplicación debido a que en la etapa de entrenamiento la red ha aprendido unos pesos capaces de generalizar para nuevos usuarios.

Estas arquitecturas también han sido utilizadas para verificación de firma, pero no con RNN, sino con TDNNs (Time Delay Neural Networks), una variación de redes convolucionales unidimensionales aplicadas a series temporales [28].

Por este motivo, el sistema propuesto en este Trabajo Fin de Grado es una arquitectura siamesa RNN que minimiza una función de coste, ver Fig. 4.1. En este sistema tenemos una red con dos firmas como entradas y cuya salida va a ser '0' si la firma realizada por el firmante es impostora o '1' si es genuina.

En concreto, el sistema propuesto en este trabajo es el que se muestra en la Fig 4.2. Se trata de un sistema basado en el uso de RNN bidireccionales (i.e. BGRU y BLSTM) con una arquitectura siamesa, con un total de dos capas RRN ocultas (*RNN hidden layers*) y una última

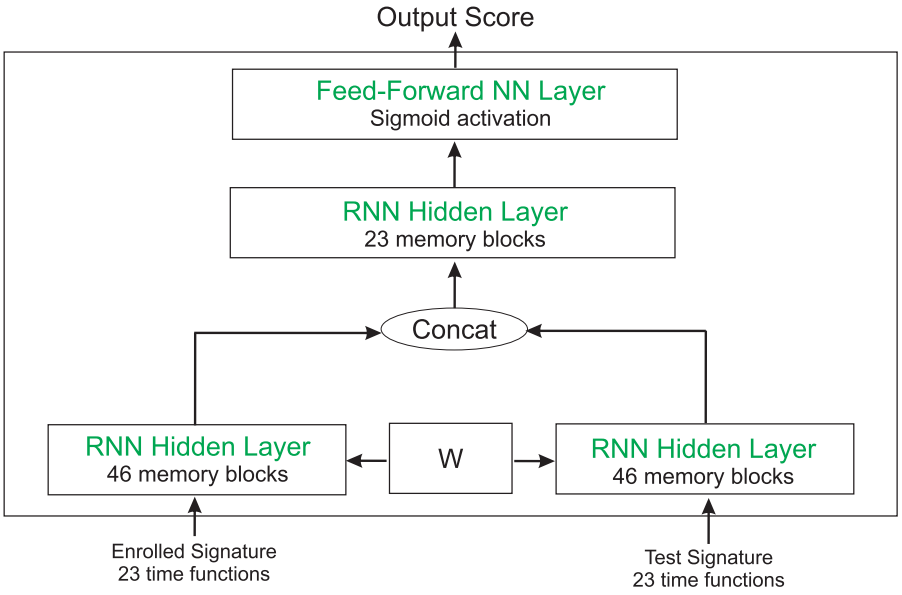


Figura 4.2: Sistema end-to-end de verificación de firma on-line propuesto. Fuente: [22].

capa *feed-forward*, explicadas a continuación:

1. **Primera capa**: formada a su vez por dos capas ocultas de tipo BGRU o BLSTM con 46 bloques de memoria cada una. La entrada de cada una de estas dos capas va a ser el conjunto de 23 funciones temporales de cada par de firmas. Las dos capas comparten los pesos entre ellas.

2. **Segunda capa**: corresponde a una capa BGRU o BLSTM con 23 bloques de memoria. La entrada de esta capa va a estar formada por las salidas de las dos capas anteriores concatenadas.

3. **Tercera capa**: formada por una capa Feed-Forward NN (*Feed-Forward Neural Network*), lo que indica que la información viaja solamente hacia delante. Esta capa consta de una función de activación sigmoide que proporciona un valor de salida entre 0 y 1 para cada pareja de firmas de entrada al sistema, donde 0 corresponde a pares de firmas genuina-impostora y 1 a pares de firmas genuina-genuina.

5

Desarrollo experimental

5.1. Dataset 1: Stylus

5.1.1. Protocolo experimental

Se han considerado como baseline dos sistemas diferentes. El primero de ellos es el caso tradicional DTW. El segundo se trata de una aproximación novedosa basada en RNN entrenada solamente con las firmas on-line de los usuarios de BiosedureID [22]. En esta base de datos las firmas fueron capturadas con un solo dispositivo y con el útil de escritura stylus, por lo que únicamente se tiene en cuenta el escenario de multi-sesión. El primer dataset propuesto, por lo tanto, va a consistir en las firmas de BiDA MDI-Sign realizadas con stylus sobre las tablets Wacom, como se muestra en la Tabla 3.3, para estudiar la habilidad de los sistemas basados en RNN para extraer características más robustas ante distintos dispositivos y escenarios.

En este dataset se van a utilizar los esquemas bidireccionales de las dos arquitecturas explicadas en la Sección 2.4.3 (i.e. BGRU, BLSTM) [1]. El proceso va a constar de tres etapas (i.e. entrenamiento, validación y evaluación). Se van a destinar un total de 867 usuarios para la etapa de entrenamiento (i.e. 80 % de los usuarios de *development*), 217 para la validación (i.e. 20 % de los usuarios de *development*) y 442 para la evaluación (i.e. 100 % de los usuarios de *evaluation*), contando con un total de 33.652, 8.424 y 15.766 firmas, respectivamente.

El entrenamiento va a consistir en introducir pares de firmas a la red junto con las etiquetas (i.e. 0 si es falsificación, 1 si es genuina). Para el entrenamiento se tienen un total de 243.664 pares de firmas, mientras que para la validación se han obtenido 60.956 pares. El método seguido para la obtención de las parejas de firmas consiste en comparar un número de firmas genuinas de *train* con un conjunto de firmas genuinas de *test* y un conjunto de falsificaciones, ver Tabla 5.1. Se han considerado las firmas de la primera sesión como firmas genuinas de train y las del resto de sesiones como firmas genuinas de test, con el objetivo de tener en cuenta el problema de **variabilidad inter-sesión**. Esto no va a ocurrir para la base de datos MCYT, en la que sólo se capturaron firmas en una sesión y se han considerado las 5 primeras como firmas genuinas de train y las 20 restantes como firmas genuinas de test. Para las parejas de entrenamiento se han considerado el mismo número de pares de firmas genuina-genuina que de genuina-impostora

[1] Las consideraciones generales para la implementación del sistema propuesto se definen en el Anexo A

	# Firmas Genuinas Train (1ª Sesión)	# Firmas Genuinas Test (Otras sesiones)	# Falsificaciones Skilled	# Comparaciones Totales
MCYT (230 usuarios)	5 (primeras)	20 (últimas)	20 (últimas)	46.000
BiosecurID (268 usuarios)	4	12	12	25.728
BiosecureDS2 (510 usuarios)	15	15	15 (últimas)	229.500
e-BioSign DS1 (30 usuarios)	4	4	4 (últimas)	1.920
e-BioSign DS2-DS3 (46 usuarios)	4	4	4 (últimas)	1.472
	# COMPARACIONES TOTALES DEVELOPMENT – DATASET 1			**304.620**

Cuadro 5.1: Distribución del número de parejas de firmas para el conjunto de Development del dataset 1.

con el objetivo de que la red aprenda de manera equitativa de los dos escenarios. En la etapa de evaluación, en cambio, se van a tener en cuenta prácticamente todas las firmas de cada usuario, excepto en el caso de BiosecureDS2, donde se han cogido menor número de firmas genuinas de *train* para no tener un número muy elevado de comparaciones de esta base de datos, ver Tabla 5.2. Se han obtenido un total de 81.372 comparaciones.

En este trabajo en concreto se va a hacer uso de un tipo de falsificación de firmas, las *skilled forgeries* (falsificaciones intencionadas), ya que es el escenario más complicado de la firma on-line. En primer lugar se va a entrenar la red utilizando las firmas genuinas y las *skilled forgeries* de cada usuario. El proceso de evaluación seguirá la misma dinámica, obteniendo los scores del sistema propuesto para las comparaciones de firmas genuina-genuina y genuina-*skilled forgerie*. Esto va a permitir se pueda realizar un análisis de los casos 1vs1 y 4vs1, explicados a continuación:

- **1vs1**: se obtendrán los scores del sistema para cada pareja de firmas considerada.
- **4vs1**: se realizará una media de los scores obtenidos de las comparaciones de cada firma genuina de test con las 4 primeras firmas genuinas de *train*. El proceso es equivalente para las *skilled forgeries*.

5.1.2. Evaluación de los resultados

Antes de comenzar a estudiar los experimentos llevados a cabo, se han obtenido los resultados de evaluar el modelo utilizado en [22] con las diferentes bases de datos que componen BiDA MDI-Sign, con el objetivo de partir de unos valores de rendimiento iniciales. Este modelo fue sólo entrenado con las firmas de BiosecurID. También se ha observado que en el estado del arte se han conseguido buenos resultados utilizando algoritmos DTW para sistemas de verificación de firma on-line, especialmente para el escenario de *random forgeries*, por lo que también se va a evaluar este primer dataset con DTW. Por ello, se parte de dos sistemas baseline (i.e DTW para BiDA MDI-Sign y RNN entrenadas sólo con BiosecurID) y se obtienen los resultados para el sistema propuesto en la Sección 4 utilizando la base de datos BiDA MDI-Sign, tanto con BGRU como con BLSTM. En la Tabla 5.3 se puede observar una comparativa del rendimiento de los cuatro sistemas para cada una de las 5 bases de datos que componen BiDA MDI-Sign, para el escenario de evaluar con *skilled forgeries*.

	# Firmas Genuinas Train (1ª Sesión)	# Firmas Genuinas Test (Otras sesiones)	# Falsificaciones Skilled	# Comparaciones Totales
MCYT (100 usuarios)	10 (primeras)	15 (últimas)	25	40.000
BiosecurID (132 usuarios)	4	12	12	12.672
BiosecureDS2 (140 usuarios)	5 (primeras)	15	20	24.500
e-BioSign DS1 (35 usuarios)	4	4	6	2.800
e-BioSign DS2-DS3 (35 usuarios)	4	4	6	1.400
	# COMPARACIONES TOTALES EVALUATION – DATASET 1			**81.372**

Cuadro 5.2: Distribución del número de parejas de firmas para el conjunto de Evaluation del dataset 1.

			MCYT	BiosecurID	BiosecureDS2	e-BioSign DS1	e-BioSign DS2-DS3
Sistemas Baseline	**DTW**	**1vs1**	7.24	6.64	12.14	11.87	8.39
		4vs1	**6.19**	5.24	10.38	9.29	8.57
	BLSTM (BiosecurID)	**1vs1**	12.30	5.60	9.32	13.21	6.31
		4vs1	13.12	4.75	8.25	11.91	3.81
Sistemas Propuestos	**BRGU**	**1vs1**	9.69	3.91	7.84	9.82	3.87
		4vs1	10.19	3.41	**7.12**	**8.21**	2.86
	BLSTM	**1vs1**	9.57	3.93	8.84	10.54	3.75
		4vs1	10	**3.35**	7.62	9.29	**2.14**

Cuadro 5.3: Rendimiento de los sistemas bajo estudio en términos de EER(%) para el escenario de entrenamiento y evaluación con *skilled forgeries*.

Se puede observar que el sistema basado en RRN propuesto junto con el aumento de datos de entrenamiento debido a la nueva base de datos BiDA MDI-Sign, han conseguido mejorar los resultados del sistema baseline entrenado sólo con BioseurID para todas las bases de datos consideradas. El sistema propuesto ha conseguido mejorar también los resultados del algoritmo DTW para todas las bases de datos menos MCYT. Tanto la arquitectura BGRU como la BLSTM obtienen unos resultados muy parecidos por lo que a partir de este experimento se va a hacer uso únicamente de la primera de ellas, ya que es más óptima y rápida de entrenar al tener dos puertas que controlan el flujo de información del bloque en lugar de las tres que tiene BLSTM. Un aspecto a destacar es la baja tasa de EER obtenida para las bases de datos BiosecurID y e-BioSign DS2-DS3, con un **3.35 %** y un **2.14 %** respectivamente. Si comparamos este resultado con los dos sistemas baseline, en BiosecurID se obtienen mejoras de **1.89 %** respecto al sistema DTW y de **1.4 %** respecto al sistema BLSTM entrenado sólo con BiosecurID, mientras que en e-BioSign DS2-DS3 se obtienen mejoras de **6.43 %** y de **1.67 %**, respectivamente. Con estos resultados se deduce que la red es capaz de aprender mejores *features* debido a que tiene más datos de entrenamiento. Otro punto importante que se puede extraer de la tabla de resultados es que el escenario de 4vs1 tiende a mejorar en torno a un **1 %** los resultados obtenidos para 1vs1 excepto para la base de datos MCYT. Esto puede deberse al protocolo de adquisición utilizado, ya que los usuarios tuvieron que realizar hasta 25 firmas en una misma sesión. En las

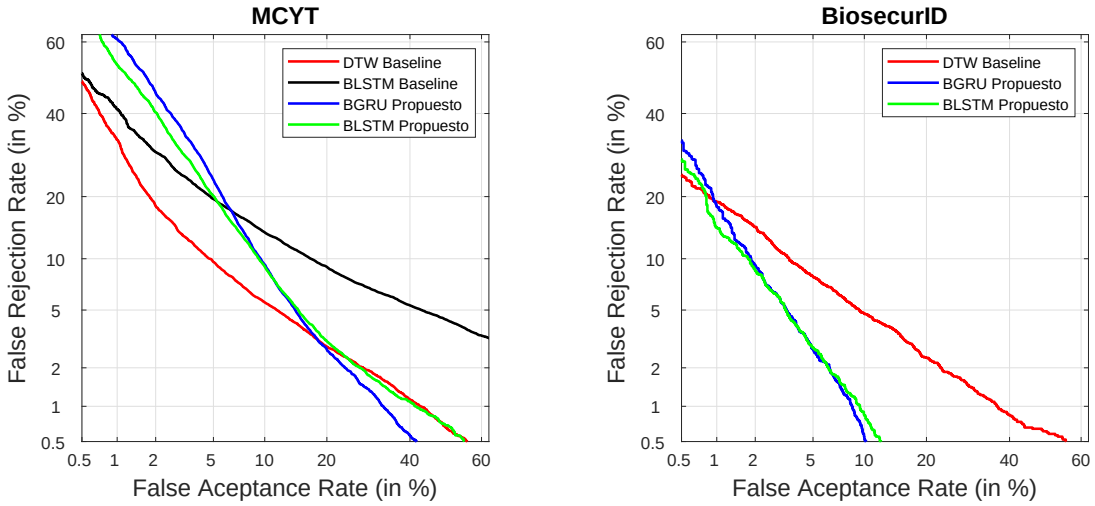


Figura 5.1: Rendimiento de los sistemas baseline y los sistemas propuestos para el dataset 1: Stylus. Escenario: 1vs1 para entrenamiento con *skilled forgeries*.

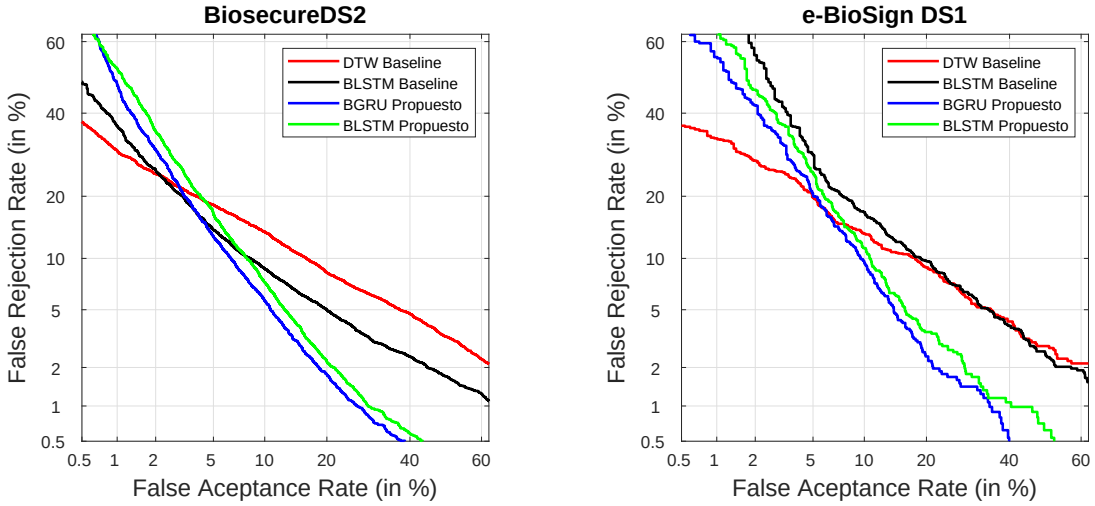


Figura 5.2: Rendimiento de los sistemas baseline y los sistemas propuestos para el dataset 1: Stylus. Escenario: 1vs1 para entrenamiento con *skilled forgeries*.

Fig. 5.1, 5.2, 5.3 se han reunido un conjunto de curvas DET en la que se observa la mejora de los sistemas propuestos respecto a los sistemas baseline para las cinco bases de datos que forman BiDA MDI-Sign.

Se ha analizado el rendimiento del sistema en función de la complejidad del usuario, habiendo clasificado a los usuarios en 3 grupos de complejidades siguiendo los trabajos recientes en el estado del arte [13], ver Fig. 5.4. Para realizar este estudio, se han escogido una serie de usuarios de cada complejidad de cada base de datos, analizando los scores obtenidos por la red RNN entrenada y observando que el comportamiento del sistema propuesto es muy similar para todas las bases de datos en cuanto a la complejidad de usuarios se refiere:

Para usuarios de **complejidad media y baja**, se ha observado que la mayor parte de los errores se producen para el caso de pares de firmas genuinas. Esto origina como resultado un sistema con un elevado valor de FRR. Sin embargo, el sistema propuesto es muy robusto para este tipo de usuarios frente a *skilled forgeries*, con bajas tasas de FAR. La Fig. 5.5 muestra algunos ejemplos de errores y aciertos obtenidos por el sistema propuesto para usuarios de baja/media

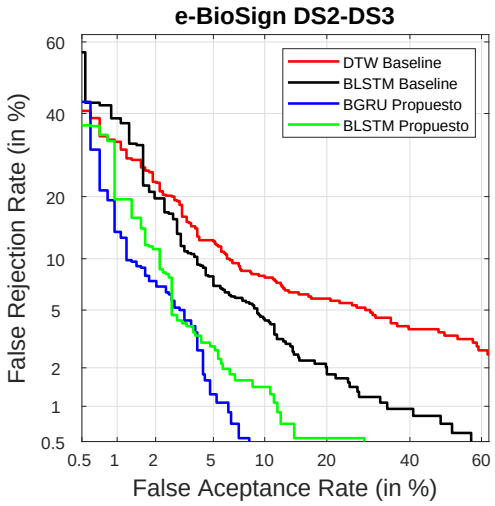


Figura 5.3: Rendimiento de los sistemas baseline y los sistemas propuestos para el dataset 1: Stylus. Escenario: 1vs1 para entrenamiento con *skilled forgeries*.

Figura 5.4: Firmas de baja, media y alta complejidad, de izquierda a derecha.

complejidad.

Para usuarios con firmas de **alta complejidad** ocurre lo contrario. La mayoría de los errores del sistema se deben a las *skilled forgeries*. Esto origina un valor elevado de FAR. Por otro lado, hay muy pocas firmas genuinas de test que esté detectando el sistema como impostoras, es decir, la tasa FRR es baja. La Fig. 5.6 muestra algunos ejemplos de errores y aciertos obtenidos por el sistema propuesto para usuarios de complejidad alta.

5.2. Dataset 2: Dedo

5.2.1. Protocolo experimental

Una vez estudiado el caso de firmas capturadas por stylus, se van a estudiar el rendimiento de nuestra arquitectura de red para las firmas realizadas con el dedo. En este caso el conjunto de firmas es mucho más reducido que en el Dataset 1, ya que solamente hay firmas realizadas con el dedo en las bases de datos e-BioSign DS1 (dispositivos W4 y W5) y e-BioSign DS2-DS3 (dispositivos W5 y W6). Ya que los resultados obtenidos por las dos arquitecturas propuestas (i.e. BGRU y BLSTM) en el Dataset 1 son muy parecidos, en este Dataset 2 se va a hacer uso de la arquitectura siamesa BGRU para llevar a cabo el análisis de las firmas de dedo, ya que es más óptima y rápida de entrenar al tener dos puertas que controlan el flujo de información del bloque en lugar de las tres que tiene BLSTM [2]. Se van a utilizar 61 usuarios para el entrenamiento (i.e. 80 % de los usuarios de *development* de dedo), 15 para validación (i.e. 20 % de los usuarios de

[2]Las consideraciones generales para la implementación del sistema propuesto se definen en el Anexo A

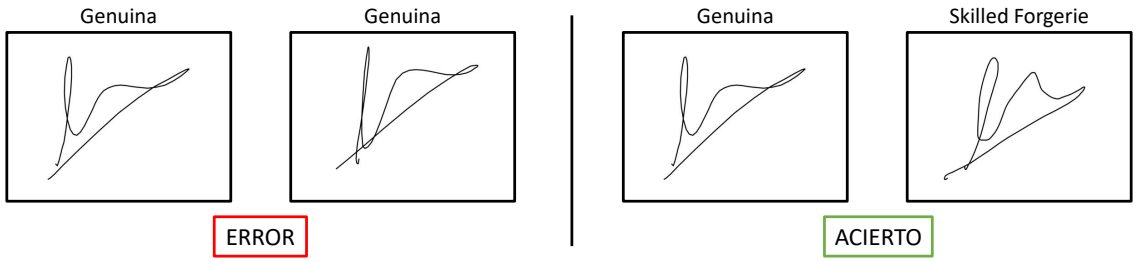


Figura 5.5: Ejemplos de error y acierto obtenidos por el sistema propuesto para usuarios de baja/media complejidad.

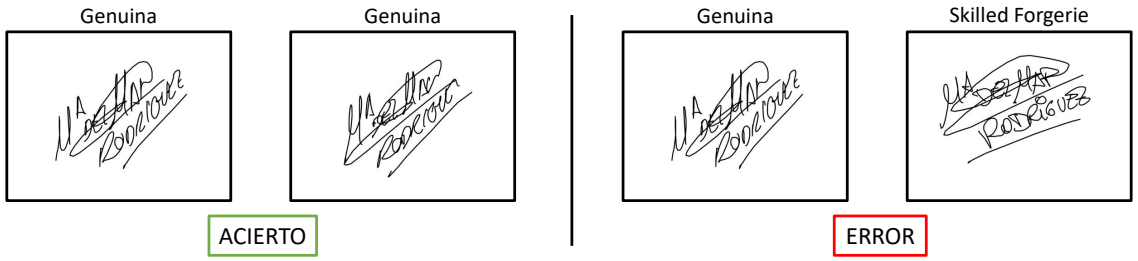


Figura 5.6: Ejemplos de error y acierto obtenidos por el sistema propuesto para usuarios de alta complejidad

development de dedo) y 70 usuarios para evaluación (i.e. 100 % de los usuarios de *evaluation* de dedo), contando con un total de 1.464, 360 y 1.960 firmas, respectivamente.

El entrenamiento va a seguir la misma dinámica que en el caso de stylus. Se van a introducir pares de firmas a la red junto con las etiquetas (i.e. 0 si es falsificación, 1 si es genuina). Para el entrenamiento se va a disponer de un total de 3.904 pares de firmas, mientras que para la validación se han obtenido 960 pares. El método de obtención de parejas es idéntico al caso de stylus, comparando un número determinado de firmas genuinas de *train* con un conjunto de firmas genuinas de *test* y un conjunto de falsificaciones, ver Tabla 5.4. Las firmas genuinas de train son las obtenidas en la primera sesión, mientras que las de test son las del resto de sesiones. Con esto se pretende tener el cuenta el problema de **variabilidad inter-sesión**. Al realizar las parejas de entrenamiento, se han considerado el mismo número de pares de firmas genuina-genuina que de genuina-impostora para que la red aprenda de manera equitativa de los dos escenarios. En la etapa de evaluación se van a tener en cuenta todas las firmas de cada usuario, obteniendo un total de 5600 pares de firmas, ver Tabla 5.5.

	# Firmas Genuinas Train (1ª Sesión)	# Firmas Genuinas Test (Otras sesiones)	# Falsificaciones Skilled	# Comparaciones Totales
e-BioSign DS1 (30 usuarios)	4	4	4 (últimas)	1.920
e-BioSign DS2-DS3 (46 usuarios)	4	4	4 (últimas)	2.944
	# COMPARACIONES TOTALES DEVELOPMENT – DATASET 2			**4.864**

Cuadro 5.4: Distribución del número de parejas de firmas para el conjunto de Development del dataset 2.

	# Firmas Genuinas Train (1ª Sesión)	# Firmas Genuinas Test (Otras sesiones)	# Falsificaciones Skilled	# Comparaciones Totales
e-BioSign DS1 (35 usuarios)	4	4	6	2.800
e-BioSign DS2-DS3 (35 usuarios)	4	4	6	2.800
	# COMPARACIONES TOTALES EVALUATION – DATASET 1			**5.600**

Cuadro 5.5: Distribución del número de parejas de firmas para el conjunto de Evaluation del dataset 2.

Al igual que en el estudio anterior, se van a utilizar un tipo de firma falsificada (i.e. *skilled forgeries*). El sistema será entrenado y evaluado en el escenario de *skilled forgeries*. Para que los resultados sean comparables a los del Dataset 1, los rendimientos del sistema se estudiarán en términos de EER(%) tanto para el escenario 1vs1 como para 4vs1, explicados en la Sección 5.1.1.

Para estudiar el rendimiento del sistema para firmas realizadas con el dedo se proponen tres métodos:

1. **Método 1**: evaluar el sistema entrenado con stylus con las firmas de dedo.
2. **Método 2**: reentrenar el modelo de stylus con las firmas de dedo.
3. **Método 3**: entrenar los pesos del sistema desde cero utilizando sólo las firmas con el dedo.

En el primer método no se ha realizado ninguna modificación de la arquitectura de la red, ya que simplemente se trata de evaluar las firmas de dedo sobre el modelo de *stylus* entrenado. En cambio, en el segundo modelo, en el que se reentrenan los pesos de la red anterior con las firmas de dedo, se han realizado varias pruebas en la forma de cargar los pesos de stylus y en la arquitectura de la red con el objetivo de mejorar el rendimiento del sistema. Estas modificaciones se han realizado en el script de python donde se define la red y el entrenamiento, y se explican en el Anexo B. De todas las pruebas realizadas, los mejores resultados se han obtenido para la última, es decir, utilizando una red idéntica a la anterior pero sustituyendo los dropouts previos a cada bloque BGRU por un dropout y un recurrent_dropout pasados como parámetros en la definición de cada bloque. Por ello, esta arquitectura se ha utilizado también para el tercer método en el que se entrena la red desde cero con el dedo.

5.2.2. Evaluación de los resultados

El escenario de firma con dedo es un problema que no ha sido todavía estudiado en profundidad por lo que, a diferencia del escenario de stylus, en este dataset no se parte de ningún sistema baseline que haya sido entrenado con este tipo de firma. Por este motivo, se ha decidido analizar el comportamiento del modelo de stylus con las firmas de dedo para partir de unos valores de rendimiento iniciales. Además, se va a utilizar un sistema DTW para evaluar el conjunto de firmas de dedo y poder comparar con el resto de resultados. En la Tabla 5.6 se muestra el rendimiento en términos de EER (%) conseguido mediante los 3 métodos propuestos en el protocolo experimental para los conjuntos de evaluación de los datasets 1 y 2. En la tabla se puede observar como el modelo entrenado con stylus no generaliza bien para otros útiles de escritura, consiguiendo un 18.75 % de tasa de EER para las firmas de dedo. En cambio, si se

	Método 1		Método 2		Método 3	
Conjunto de Entrenamiento	**Stylus (DS1)**		**Stylus+Dedo (DS1+DS2)**		**Dedo (DS2)**	
Conjunto de Evaluación	**Stylus (DS1)**	**Dedo (DS2)**	**Stylus (DS1)**	**Dedo (DS2)**	**Stylus (DS1)**	**Dedo (DS2)**
EER (%)	**8.34**	18.75	12.75	**13.93**	37.54	25.85

Cuadro 5.6: Rendimiento del sistema bajo estudio en términos de EER(%) para los los 3 métodos definidos para todo el conjunto de evaluación de los datasets 1 y 2.

reentrena el modelo de stylus con las firmas de dedo, se consigue mejorar el rendimiento para las firmas de dedo en un **4.82 %**, pero se empeora para firmas realizadas con stylus en un 4.41 %, consiguiendo un modelo más equilibrado para distintos útiles de escritura pero con peor rendimiento en firmas realizadas con stylus que el anterior. Por último, si entrenamos los pesos del sistema desde cero con las firmas de dedo, no se consiguen buenos resultados ni para stylus ni para dedo, teniendo una tasa de EER de 25.85 % para el último caso, debido al menor número de firmas disponibles para este escenario. Comparando estos últimos resultados con los del modelo entrenado con stylus, se puede deducir que la arquitectura siamesa propuesta necesita un tamaño de datos grande para poder aprender las mejores *features* posibles y ser robusta a distintos útiles de escritura y distintas sesiones de captura de firmas. El modelo reentrenado con firmas de dedo mejora los resultados de EER para firmas de dedo en un **11.92 %** respecto al modelo entrenado únicamente con dedo. Esto se debe a que se ha hecho uso de un tamaño de datos de entrenamiento mayor y de que se parte de un sistema con unos pesos previamente entrenados para el problema de firma, en vez de tener que entrenar los pesos de la red desde cero (método 3).

Para un mayor análisis de la influencia de las firmas de dedo en el entrenamiento del sistema propuesto, se va a analizar a fondo el rendimiento en términos de EER (%) de los métodos 2 y 3 en comparación con el algoritmo DTW utilizado como baseline, ya que en el método 1 no se utilizan las firmas de dedo en el entrenamiento de los pesos del sistema. En la Tabla 5.7 se recogen los resultados obtenidos para cada una de las dos bases de datos con firmas de dedo, para el caso de entrenar y evaluar con *skilled forgeries*. Como ya se ha podido deducir de la tabla anterior con el rendimiento de los diferentes métodos, el entrenamiento desde cero de los pesos del sistema con las firmas de dedo no consigue obtener un buen rendimiento para ninguna de las dos bases de datos. Cabe destacar el **8.21 %** de EER obtenido por el sistema propuesto

			e-BioSign DS1	e-BioSign DS2-DS3
Sistema Baseline	**DTW**	**1vs1**	23.54	14.64
		4vs1	20.36	11.07
Sistemas Propuestos	**BGRU (Método 2)**	**1vs1**	20.09	8.66
		4vs1	**17.74**	**8.21**
	BGRU (Método 3)	**1vs1**	31.25	19.61
		4vs1	30	20

Cuadro 5.7: Rendimiento de los sistemas bajo estudio en términos de EER(%) para el escenario de entrenamiento y evaluación con *skilled forgeries*.

reentrenado desde los pesos del modelo de stylus para la base de datos e-BioSign DS2-DS3, que mejora en un **2.86 %** el mejor resultado de DTW y en un **11.4 %** el mejor resultado de aplicar el método 3. Con este resultado se puede deducir que el sistema propuesto aprende mejores *features* cuanto mayor es el número de datos de entrenamiento, y se adapta mejor al escenario de interoperabilidad de útiles de escritura cuando se entrena con firmas realizadas tanto con dedo como con stylus.

En la Fig. 5.7 se muestran un conjunto de curvas DET en las que se puede apreciar cómo el hecho de reentrenar los pesos del modelo de stylus mejora el sistema baseline DTW y cómo el hecho de entrenar desde cero con muchas menos firmas que en el primer dataset empeora los resultados.

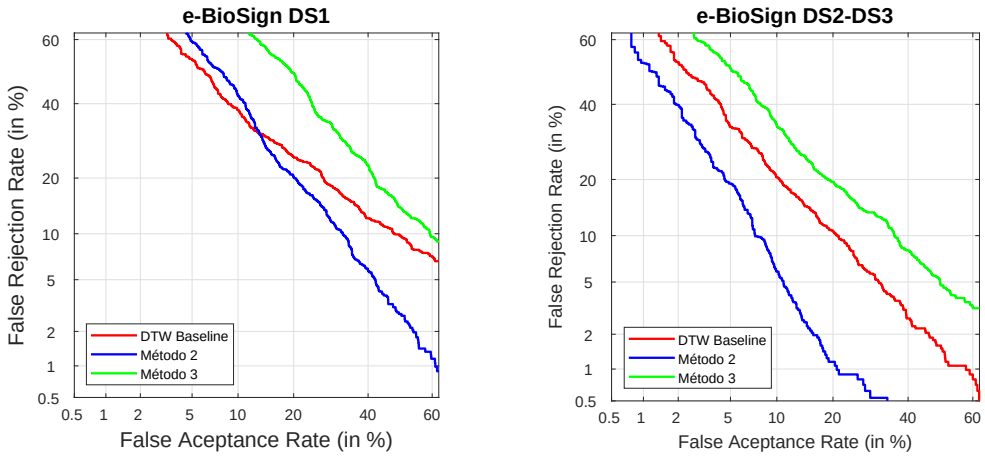


Figura 5.7: Rendimiento del sistema baseline y los métodos 2 y 3 para cada base de datos del dataset 2: Dedo. Escenario: 1vs1.

6

Conclusiones y trabajo futuro

La primera etapa del presente Trabajo Fin de Grado se ha centrado en la creación de una nueva base de datos que contemple los escenarios de adquisición actuales y con un volumen de usuarios no visto hasta el momento en ninguna base de datos de firma on-line existente. Para ello, se han utilizado firmas capturadas con 3 tablets Wacom diferentes, dos tipos de tablet Samsung y un smartphone, en diferentes sesiones, permitiendo así estudiar el rendimiento de un sistema de verificación de firma on-line en escenarios de interoperabilidad de dispositivos de captura y útiles de escritura, así como variabilidad inter-sesión. Dentro de la base de datos creada en el presente Trabajo Fin de Grado, se han tenido en cuenta dos datasets para los experimentos (i.e. stylus y dedo). Los resultados con firmas de dedo han mejorado los resultados obtenidos hasta el momento utilizando sistemas tradicionales como DTW con un 2.86 % de mejora de EER. Para las firmas realizadas con stylus se ha conseguido mejorar los resultados del estado del arte para casi todas las bases de datos [22], permitiendo tener un sistema de verificación de firma capaz de generalizar entre diferentes sesiones y dispositivos de captura, debido a la gran cantidad de datos utilizados en el entrenamiento. Por último, se ha realizado un análisis del sistema propuesto, concluyendo que actúa de manera diferente en función de la complejidad de la firma, siendo muy robusto frente a ataques en usuarios de baja y media complejidad y más débil en usuarios de alta complejidad.

A pesar de los buenos resultados obtenidos, existe un importante trabajo por hacer con las firmas de la base de datos BiDA MDI-Sign. Algunos estudios que se podrían llevar a cabo en un futuro son los siguientes:

- Proponer distintos sistemas RRN adaptados a las complejidades de los usuarios.
- Proponer el uso de otras arquitecturas más robustas que consideren mayor número de firmas del usuario y no sólo pares de firmas.
- Utilizar modelos pre-entrenados en otros escenarios con mayores volúmenes de datos (e.g. reconocimiento de escritura on-line) y adaptarlo al problema de firma.
- Analizar el resto de datasets creados en BiDA MDI-Sign.
- Realizar un estudio de las complejidades de las firmas realizadas con el dedo.
- Utilizar la red neuronal como extractor de características para entrenar un sistema basado en SVM.
- Entrenar la red con *random forgeries* o con una mezcla de *skilled* y *random*.

Glosario de acrónimos

- **RNN**: *Recurrent Neural Networks*
- **EER**: *Equal Error Rate*
- **FA**: Falsa Aceptación
- **FR**: Falso Rechazo
- **FAR**: *False Acceptance Rate*
- **FRR**: *False Rejection Rate*
- **DET**: *Detection Error Trade-off*
- **NN**: *Neural Networks*
- **CNN**: *Convolutional Neural Networks*
- **HMM**: *Hidden Markov Models*
- **DTW**: *Dynamic Time Warping*
- **SVM**: *Support Vector Machines*
- **DCT**: *Discret Cosine Transform*
- **LSTM**: *Long Short-Term Memory*
- **GRU**: *Gated Recurrent Unit*
- **BLSTM**: *Bidirectional LSTM*
- **BGRU**: *Bidirectional GRU*
- **BRNN**: *Bidirectional RNN*
- **TDNN**: *Time Delay Neural Networks*

Bibliografía

[1] J. Fierrez, A. Pozo, M. Martinez-Diaz, J. Galbally, and A. Morales. Benchmarking touchscreen biometrics for mobile authentication. *IEEE Trans. on Information Forensics and Security*, 13(11):2720–2733, November 2018.

[2] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. Incorporating touch biometrics to mobile one-time passwords: Exploration of digits. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPR-W*, June 2018.

[3] R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia, A. Acien, and R. Tolosana. e-biosign tool: Towards scientific assessment of dynamic signatures under forensic conditions. In *Proc. IEEE BTAS*, September 2015.

[4] A. Ross, K. Nandakumar, and A. K. Jain. *Handbook of multibiometrics: human recognition systems; 1.* Springer, Dordrecht, 2006.

[5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[6] J. Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.

[7] G. E. Hinton. Boltzmann machine. *Scholarpedia*, 2(5):1668, 2007.

[8] S. Lawrence, C.L. Giles, Ah Chung Tsoi, and A.D. Back. Face recognition: a convolutional neural-network approach. *Neural Networks, IEEE Transactions on*, 8(1):98–113, January 1997.

[9] O. Costilla Reyes, R. Vera-Rodriguez, P. Scully, and K. B. Ozanyan. Analysis of spatio-temporal representations for robust footstep recognition with deep residual neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (99), 2018.

[10] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber. Convolutional neural network committees for handwritten character classification. In *2011 International Conference on Document Analysis and Recognition*, pages 1135–1139, 2011.

[11] A. Graves and N. Jaitly. Towards end-to-end speech recognition with recurrent neural networks. In *Proc. ICML*, pages 1764–1772. JMLR.org, 2014.

[12] A. Petrosian, D. Prokhorov, R. Homan, R. Dasheiff, and D. Wunsch. Recurrent neural network based prediction of epileptic seizures in intra- and extracranial eeg. *Neurocomputing*, 30(1):201 – 218, 2000.

[13] R. Tolosana, R. Vera-Rodriguez, R. Guest, J. Fierrez, and J. Ortega-Garcia. Complexity-based biometric signature verification. In *14th IAPR International Conference on Document Analysis and Recognition*, 2017.

[14] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally. Mobile signature verification: Feature robustness and performance comparison. *IET Biometrics*, 3(4):267–277, December 2014.

[15] N. Houmani, S. Garcia-Salicetti, B. Dorizzi, and *et al.* Biosecure signature evaluation campaign (esra'2011): evaluating systems on quality-based categories of skilled forgeries. In *Proc. IJCB*, pages 1–10, Oct 2011.

[16] M. Martinez-Diaz and J. Fierrez. *Signature Databases and Evaluation*, pages 1367–1375. Springer, 2015. ISBN 978-1-4899-7487-7, re-edited from 2009.

[17] J. Fierrez, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez. Hmm-based on-line signature verification: feature extraction and signature modeling. *Pattern Recognition Letters*, 28(16):2325–2334, December 2007.

[18] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia. Benchmarking desktop and mobile handwriting across cots devices: the e-biosign biometric database. *PLOS ONE*, 5(12), 2017.

[19] Y. Liu, Z. Yang, and L. Yang. Online signature verification based on dct and sparse representation. *IEEE Transactions on Cybernetics*, 45(11):2498–2511, Nov 2015.

[20] A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern Recogn.*, 38(12):2270–2285, December 2005.

[21] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, November 1997.

[22] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. Exploring recurrent neural networks for on-line handwritten signature biometrics. *IEEE Access*, pages 1 – 11, 2018.

[23] S. Dey, A. Dutta, J. I. Toledo, S. K. Ghosh, J. Lladós, and U. Pal. Signet: Convolutional siamese network for writer independent offline signature verification. *ArXiv*, 2017.

[24] S. Lai, L. Jin, and W. Yang. Online signature verification using recurrent neural network and length-normalized path signature. *ICDAR*, 2017.

[25] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez, V. Espinosa, A. Satue, I. Hernaez, J. J. Igarza, C. Vivaracho, D. Escudero, and Q. I. Moro. Mcyt baseline corpus: A bimodal biometric database. *IEE Proceedings Vision, Image and Signal Processing, Special Issue on Biometrics on the Internet*, 150(6):395–401, December 2003.

[26] J. Fierrez, J. Galbally, J. Ortega-Garcia, M. R. Freire, F. Alonso-Fernandez, and D. Ramos *et al.* Biosecurid: A multimodal biometric database. *Pattern Analysis and Applications*, 13(2):235–246, May 2010.

[27] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, and *et al.* The multi-scenario multi-environment biosecure multimodal database (bmdb). *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(6):1097–1111, June 2010.

[28] J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Säckinger, and R. Shah. Signature verification using a siamese time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(4):669–688, 1993.

A

Consideraciones generales de implementación de los sistemas propuestos

Para la experimentación realizada se han utilizado la herramienta **Keras** sobre **TensorFlow** y sobre **Theano**. En primer lugar, cabe destacar que para entrenar y evaluar la red propuesta en el apartado anterior, las entradas tienen que tener el mismo tamaño. Pero en todas las bases de datos utilizadas cada firma tiene un número de muestras independiente. Por este motivo, se ha obtenido un histograma de los tamaños de todas las firmas de Megasign, en número de muestras, tanto de entrenamiento como de evaluación. Una vez obtenido el histograma, se ha establecido un número común de muestras para todas las firmas, teniendo en cuenta que con un número demasiado pequeño se va a perder mucha información y con un número demasiado grande va a aumentar mucho el coste computacional del entrenamiento de la red. El valor escogido ha sido de 2000 muestras, obteniendo para cada firma una matriz de tamaño 2000x23 con los valores de las funciones temporales calculadas, a partir de los datos de las firmas sin normalizar, como se explica en la sección anterior. Estas matrices son ficheros .mat guardados de tal manera que no se aplique ningun tipo de compresión a los datos para evitar errores en el entrenamiento del sistema.

Debido a que el sistema propuesto es una red siamesa con dos entradas, se van a introducir las firmas por parejas. Para que el sistema sepa qué parejas coger en cada momento, se han elaborado a través de Matlab ficheros .txt en los que en cada línea aparecen dos nombres (i.e. firma genuina de *train* y firma con la que se va a comparar) y la etiqueta (i.e. 1 si la segunda firma es genuina y 0 si es impostora). Se ha generado un fichero de parejas para cada escenario de evaluación y entrenamiento considerado.

Debido al gran tamaño de datos que hay que cargar e introducir en el ordenador, se ha modificado el código de python utilizado para definir y entrenar el modelo de red propuesto. En experimentos previos se cargaban los datos directamente en memoria. En este caso, se ha modificado la forma de entrenar el modelo, cambiando el comando $'fit'$ de Keras por $'fit\_generator'$, para el cual se ha diseñado una función en Python que va cargando en memoria las parejas de firmas en *batches* de 100 con el objetivo de no saturar computacionalmente el ordenador. Otro aspecto a destacar es que se ha intentado explorar el uso del *batch normalization* junto con RNN, ya que en CNN permiten obtener muy buenos resultados. Tras varios intentos e investigaciones no se ha podido implementar debido a que todavía no es una técnica muy utilizada en redes recurrentes.

B

Pruebas de código del Dataset 2: Dedo

A continuación se explican las pruebas realizadas en Keras sobre TensorFlow para conseguir el mejor rendimiento posible del sistema propuesto para el reentrenamiento de los pesos obtenidoc con stylus utilizando las firmas de dedo (i.e. método 2 del protocolo experimental):

- **Prueba 1**: en un primer intento de cargar los pesos del modelo de stylus, se ha vuelto a definir la misma red que anteriormente y se ha cargado el modelo mediante la función *'load_weights'*. Después se ha reentrenado la red.
- **Prueba 2**: se ha cargado el modelo con *'load_model'* directamente sin definir de nuevo la red y se ha reentrenado.
- **Prueba 3**: debido a que las dos pruebas anteriores no proporcionaban resultados muy convincentes, se procede a congelar ciertas capas de la red con el objetivo de no reentrenar todos los pesos aprendidos en el entrenamiento sobre stylus. En esta prueba se propone congelar el entrenamiento de las dos primeras capas bidireccionales, es decir, entrenar la última capa bidireccional y la capa *dense* del final.
- **Prueba 4**: siguiendo la filosofía del apartado anterior, se procede a congelar todas las capas bidireccionales y reentrenar únicamente la capa *dense* del final. Estas dos últimas pruebas tampoco han conseguido unos resultados convincentes.
- **Prueba 5**: en una quinta prueba, se pensó en introducir otra capa *dense* de tamaño 10 justo antes de la capa *dense* del final de la red. Sin embargo, esta modificación no ha podido ser llevada a cabo debido a que la nueva arquitectura tiene 4 capas y el modelo de stylus introduce pesos para tres capas únicamente.
- **Prueba 6**: en la arquitectura de red principal se hace uso de la herramienta *'dropout'* en la entrada de cada capa para evitar el *overfitting* (sobreajuste). Pero actualmente, en la definición de cada bloque GRU en Keras, hay un parámetro llamado *'recurrent_dropout'* que aleatoriza las conexiones entre los bloques recurrentes. En esta prueba se propone combinar estos dos parámetros de las siguientes maneras:
  - Dropout antes de los bloques BGRU y dentro de cada uno dropout y recurrent_dropout.

- Dropout antes de los bloques BGRU y dentro de cada uno recurrent_dropout únicamente.
- Sólo recurrent_dropout dentro de cada bloque BGRU.
- Dropout y recurrent_dropout dentro de cada bloque BGRU.