Universidad Autónoma de Madrid
Escuela politécnica superior

Máster Universitario en Ingeniería de Telecomunicación

Trabajo Fin de Grado

NUEVOS ESQUEMAS DE VERIFICACIÓN DE FIRMA MANUSCRITA DINÁMICA: ANÁLISIS DE LA COMPLEJIDAD Y FUSIÓN DE SISTEMAS

Autor: Miguel Caruana Montes
Tutor: Rubén Vera Rodríguez

JUNIO 2021

NUEVOS ESQUEMAS DE VERIFICACIÓN DE FIRMA MANUSCRITA DINÁMICA: ANÁLISIS DE LA COMPLEJIDAD Y FUSIÓN DE SISTEMAS

Autor: Miguel Caruana Montes
Tutor: Rubén Vera Rodríguez

Biometrics and Data Pattern Analytics - BiDA Lab
Dpto. de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Universidad Autónoma de Madrid
JUNIO 2021

Resumen

Resumen

En este Trabajo de Fin de Máster se proponen dos enfoques distintos para la mejora de sistemas de verificación de firma dinámica: el uso de la información de la complejidad de las firmas y la fusión de sistemas de verificación de firma dinámica y estática. Con este objetivo en mente, el trabajo realizado se ha divido en tres fases. Antes de comenzar esta investigación, se revisaron trabajos del estado del arte concernientes a cada uno de los tres temas desarrollados: los sistemas de verificación de firma dinámica, los sistemas de verificación de firma estática y la fusión de ambos tipos de sistemas.

En la primera parte de este trabajo, se ha realizado un análisis en profundidad de los efectos de la complejidad de las firmas en sistemas de verificación de firma dinámica. Se han considerado tres grupos de complejidad: alta, media y baja. Considerando esta división de usuarios, se ha evaluado el rendimiento de dos sistemas del estado del arte, un sistema tradicional DTW y un sistema, recientemente propuesto, basado en redes neuronales recurrentes cuyos datos de entrada están alineados temporalmente (TA-RNN). La base de datos utilizada para la evaluación del rendimiento ha sido DeepSignDB, una de las bases de datos disponibles con mayor número de usuarios. Después, se han explorado diferentes propuestas para la mejora de sistemas de verificación de firma dinámica, todas basadas en el uso de la información de complejidad. El experimento con mejores resultados ha consistido en el entrenamiento de un sistema con un número de usuarios equilibrado respecto a la complejidad de su firma. El análisis de los efectos de la complejidad también se ha llevado a cabo sobre las firmas realizadas a dedo disponibles en DeepSign.

En una segunda parte, se han estudiado diferentes aproximaciones para el desarrollo de sistemas de verificación de firma estática. Partiendo de un sistema de extracción de características del estado del arte se han explorado dos vías para mejorar el sistema, el uso de una arquitectura siamesa y el uso de la función de pérdidas triplet loss. El mejor rendimiento ha sido obtenido por un modelo que partiendo de un extractor de características basado en el sistema de referencia y entrenado con DeepSignDB realiza la comparación de las características mediante una red neuronal.

Finalmente, en la última parte de este trabajo se han evaluado los beneficios de la fusión de sistemas de verificación dinámicos y estáticos. En concreto, se ha analizado la combinación de sistemas a nivel de puntuaciones. Como resultado de la fusión, se ha obtenido un sistema que mejora los resultados del estado del arte.

Palabras Clave

Sistema Biométrico, Verificación, Firma dinámica, Firma estática, Complejidad, Fusión de Sistemas, DeepSignDB, RNN, CNN.

Abstract

In this Master's Thesis, two different approaches for the improvement of on-line signature verification systems are proposed: to use signature complexity information and the fusion of on-line and off-line signature verification systems. With this goal in mind, the research performed has been divided into three phases. Before starting this research, state-of-the-art works concerning each of the three topics developed were reviewed: on-line signature verification systems, off-line signature verification systems and the fusion of both types of systems.

In the first part of this work, an in-depth analysis on how the complexity of signatures affects the performance in on-line signature verification systems has been carried out. Three complexity groups have been considered: high, medium and low. Considering this division of users, the performance of two state-of-the-art systems, a traditional DTW system and a system based on time aligned recurrent neural networks (TA-RNN), has been evaluated. The database used for the performance evaluation has been DeepSignDB, one of the databases available with the largest number of users. Then, different proposals for the improvement of dynamic signature verification systems have been explored, all based on the use of complexity information. The experiment with the best results consisted of training a system with a balanced number of users with respect to the complexity of their signature. The analysis of the complexity effects was also carried out on the finger signatures available in DeepSign.

In a second part, different approaches for the development of off-line signature verification systems have been studied. Starting from a state-of-the-art feature extraction system, two ways to improve the system have been explored, the use of a Siamese architecture and the use of the triplet loss function. The best performance has been obtained by a model that starting from a feature extractor based on the baseline system and trained with DeepSignDB performs feature matching using a neural network.

Finally, in the last part of this work, the benefits of the fusion of on-line and off-line verification systems have been evaluated. Specifically, the combination of systems has been analyzed at the score level. As a result of the fusion, a system that improves the state-of-the-art results has been obtained.

Key words

Biometric System, Verification, On-line Signature, Off-line Signature, Complexity, Fusion of Systems, DeepSignDB, RNN, CNN.

Agradecimientos

En primer lugar, me gustaría dar las gracias a Rubén Vera, quién ha sido mi tutor durante el TFG y el TFM. Rubén ha estado siempre dispuesto a ayudarme, siempre guiándome. Gracias a él he tenido la oportunidad de publicar un *paper* en el congreso ICPR recogiendo parte del trabajo aquí descrito. Sin su inestimable ayuda el resultado de este TFM no hubiera sido el mismo.

También me gustaría dar las gracias al grupo de investigación BiDA por acogerme en su laboratorio desde la realización del TFG y por el buen trato recibido, permitiéndome tener a mi disposición para realizar experimentos hasta dos ordenadores en caso de ser necesario.

A mis compañeros de carrera gracias por hacer más amenas las clases y las prácticas, aunque no siempre aprovechásemos el tiempo al máximo; nunca está de más desconectar un poco antes de volver a sumergirse en los estudios. En concreto, me gustaría agradecerle a Andrea su paciencia a la hora de resolverme dudas, siempre dispuesta a responder mis preguntas.

Por último, dar las gracias a mi familia y amigos de toda la vida, que siempre han estado ahí para apoyarme en los buenos y malos momentos, por todos estos años en los que siempre habéis estado a mi lado.

Miguel Caruana Montes

Junio 2021

Índice general

Índice de Figuras

Índice de Tablas

1 Introducción

1.1. Motivación

En un mundo en constante digitalización los sistemas de reconocimiento biométricos aparecen como sustitutos robustos a las clásicas contraseñas alfanuméricas. Muchos de estos sistemas han obtenido una gran acogida dentro de la sociedad, como los sistemas de huella dactilar o los sistemas de reconocimiento facial.

Los sistemas de verificación de firma manuscrita consiguen adaptar una forma tradicional de autenticar la identidad de una persona al entorno digital. En concreto, los sistemas de firma dinámica aprovechan los datos temporales que podemos recopilar al firmar en dispositivos electrónicos, como una tablet, para mejorar el rendimiento de los sistemas de verificación.

Este Trabajo de Fin de Máster se encuadra como continuación del trabajo previo realizado en el Trabajo de Fin de Grado "Verificación de firma dinámica mediante redes neuronales recurrentes" [1]. Dicho trabajo estaba enfocado a mejorar sistemas de verificación de firma dinámica usando arquitecturas de Deep Learning.

Se utiliza dicho TFG como punto de partida y se continúan las líneas de trabajo propuestas: usar la información de complejidad de las firmas para mejorar el rendimiento de los sistemas de verificación de firma dinámica. Asimismo, también se plantea usar la información estática de las firmas en conjunto con la información dinámica, tratando así de incrementar la efectividad de estos sistemas.

1.2. Objetivos

Este trabajo tiene dos objetivos principales:

- **Analizar el efecto de la complejidad de las firmas en sistemas de verificación de firma dinámica.**

  Partiendo del sistema de detección de la complejidad desarrollado durante el TFG [1], se llevará a cabo un análisis del rendimiento de sistemas de verificación sobre firmas de distintas complejidades (baja, media, alta). Se estudiará no sólo sobre qué tipo de firmas

funcionan mejor estos sistemas, sino si podemos aprovechar esta información para realizar sistemas de verificación específicos por complejidad y de esta forma mejorar el rendimiento global. Este análisis se realizará tanto para firmas realizadas con stylus como para firmas realizadas con el dedo.

- **Combinar sistemas de firma estática y dinámica para mejorar el rendimiento de sistemas de verificación de firma dinámica.**

  Se desarrollará un sistema de verificación de firma estática para extraer información estática sintética de las firmas dinámicas. Posteriormente esta información se integrará junto con la información dinámica en un sistema de verificación que combine ambas aproximaciones. Esta combinación se puede realizar de varias formas. Una de ellas consiste en calcular por separado la puntuación resultante de cada sistema de verificación y luego realizar una combinación de los resultados para obtener la puntuación final. Por otro lado, se puede realizar una combinación a nivel de características, extrayendo un vector de características de cada sistema de verificación y finalmente entrenando un sistema que tenga como entrada los dos tipos de características (dinámicas y estáticas). Este trabajo se centrará únicamente en la combinación a nivel de puntuación.

En definitiva, en este trabajo se propone como objetivo la mejora de sistemas de verificación de firma dinámica mediante la incorporación de nueva información a los mismos.

2

Estado del arte

2.1. Biometría: una forma segura de identificación

La seguridad es un área que, cada vez más, se busca potenciar para evitar accesos no autorizados que puedan conllevar riesgos como el robo de información o el robo de identidad, entre otros. En los últimos tiempos hemos visto diversos ataques que han puesto de manifiesto la debilidad de los sistemas actuales. Uno de los principales vectores de ataque por el que los hackers comprometen un sistema es la debilidad de las contraseñas.

En nuestro día a día utilizamos diversos servicios en Internet como banca online, plataformas de vídeo bajo demanda o el ya estándar correo web. Esto se traduce en una amplio número de contraseñas que debemos recordar, desde las clásicas contraseñas alfanuméricas hasta los patrones de desbloqueo de los smartphones. Normalmente, para evitar olvidar este gran número de contraseñas solemos crearlas a partir de datos o experiencias personales, como la fecha de nacimiento, nombre de familiares, etc. Sin embargo, esta no es una práctica recomendable ya que hace que nuestras contraseñas sean vulnerables. Por otro lado, utilizar contraseñas largas y complejas hace que sea difícil recordarlas y nos hace más propensos a olvidarlas.

La biometría es el estudio de los rasgos físicos y/o conductuales de una persona para la identificación inequívoca de la misma. Estos rasgos reciben el nombre de rasgos biométricos y se erigen como solución a los problemas de las contraseñas tradicionales, proporcionando al usuario una contraseña fácilmente recordable y difícilmente replicable. Entre los diferentes rasgos biométricos encontramos algunos con un uso muy extendido como la huella dactilar o la cara [2, 3, 4, 5, 6, 7, 8]. Estos rasgos suelen ser muy discriminatorios pero tienen un problema, no se pueden cambiar y por lo tanto en caso de que quedase expuesta esta "*contraseña*"no podría ser reemplazada.

Por otro lado, están los rasgos conductuales como la voz, la forma de caminar o la escritura [9, 10, 11, 12, 13]. Uno de estos rasgos conductuales con una gran aceptación es la firma manuscrita [14, 15, 16]. La firma es un rasgo conductual que, a diferencia de la cara o la huella, sí es susceptible de cambio, de manera que en caso de verse comprometida sólo habría que crear una nueva. Sumado a ello, la gran tradición con la que cuenta el uso de la firma como medio de autenticación la hacen un candidato perfecto para mejorar la seguridad de las contraseñas.

2.2. Sistemas de verificación de firma manuscrita

Un sistema de verificación es una herramienta que permite comprobar si la identidad de una persona coincide con la que esta reclama, a través de la información aportada por la misma. En el caso de los sistemas de verificación de firma manuscrita la información aportada es la firma. Esta información es comparada con la información almacenada en el sistema del usuario que se dice ser y se verifica si los datos aportados son verídicos o son una falsificación. Encontramos dos tipos de falsificaciones:

- **Skilled forgeries**: Son falsificaciones de la firma original que intentan asemejarse a ella lo máximo posible. Su traducción en español sería falsificaciones intencionadas.
- **Random forgeries**: Ocurre cuando se intenta verificar la identidad de un usuario usando la firma de otro usuario distinto. La traducción al español de este término sería falsificación aleatoria.

Es importante que los sistemas de verificación sean capaces de detectar ambos tipos de falsificaciones. En general estos sistemas suelen tener un peor rendimiento frente a las skilled forgeries debido a la baja variabilidad inter-clase. Sumado a esto la alta variabilidad intra-clase y la baja estabilidad dificultan en cierta medida el desarrollo de sistemas de verificación robustos.

La estructura de los sistemas de verificación está compuesta por seis fases, como podemos ver en la Figura 2.1, las cuáles se describirán brevemente a continuación.

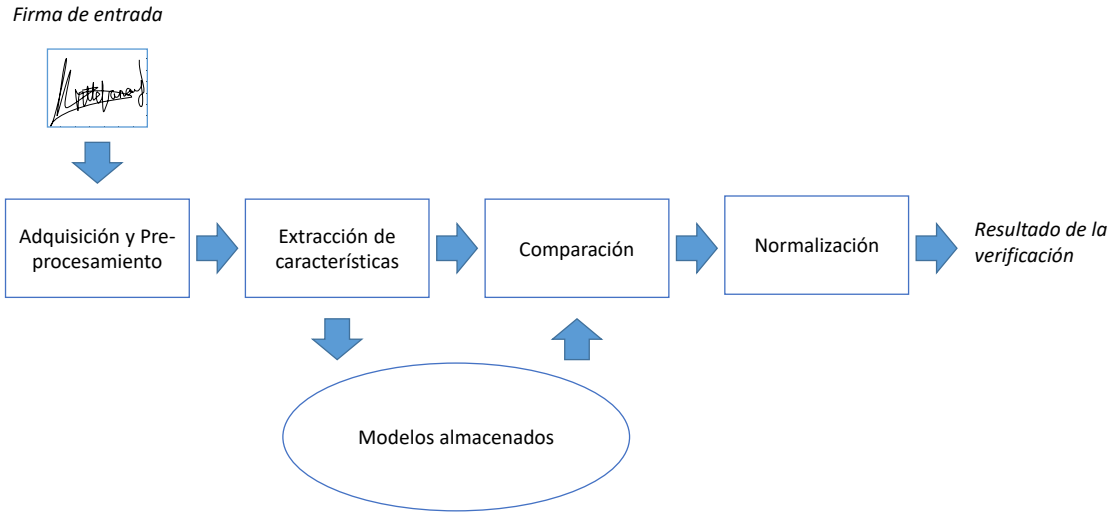


Figura 2.1: Estructura de un sistema de verificación de firma manuscrita.

- **Captura de datos**: Es la primera etapa y en ella se capturan las firmas de los distintos usuarios. Las firmas pueden ser dinámicas y en este caso se captura la información temporal del proceso de firma o estáticas, de las que solo se captura la imagen de la firma. Esta imagen se puede obtener tanto de firmas realizadas en papel como de firmas realizadas en tableta de las que previamente se ha obtenido las características temporales. A este último tipo de firmas se las denomina firmas sintéticas.
- **Pre-procesamiento**: Tras la captura de las firmas se aplican diversos procedimientos para mejorar la calidad de las señales obtenidas, como filtros para eliminar ruido o procesado de

imagen para el aumento de la resolución. Estas técnicas de procesado también se utilizan para asegurar la interoperabilidad entre distintos sistemas, dando el mismo formato a los datos extraídos por sistemas de captura diferentes.

- **Extracción de características**: Generalmente sobre los datos pre-procesados no se realiza directamente el cálculo de similitud, sino que a partir de estos se extraen una serie de características que luego sí se usaran para realizar dicho cálculo. Principalmente encontramos dos tipos de sistemas: aquellos que extraen **características globales** (e.g nº de pen-ups, velocidad de la firma) y aquellos que extraen **series temporales** (e.g presión, trayectoria).
- **Registro**: Una vez extraídas las características de las firmas estas se almacenan en bases de datos. Existen dos tipos de registro, el **basado en modelo** y el **basado en referencia**.
- **Cálculo de similitud**: Después del proceso de extracción de características se procede a realizar la comparación entre firmas. Los sistemas basados en características globales utilizan distancias como medida de similitud. Por otro lado, los sistemas basados en características temporales utilizan otro tipo de técnicas como Hidden Markov Models (HMM) [17], Support Vectors Machine (SVM) [18], Dynamic Time Warping (DTW) [19] y técnicas de Deep Learning (DL) [20] .
- **Normalización**: Finalmente, tras obtener los resultados de la etapa de comparación estos se normalizan a un rango común. Algunas de las técnicas de normalización utilizadas son *Decimal Scaling*, *Double Sigmoid* o *Biweight estimators* [21].

En la actualidad podemos encontrar dos tipos de sistemas de verificación de firma manuscrita atendiendo a al tipo de firma capturada: los sistemas de firma dinámica y los sistemas de firma estática.

2.3. Sistemas de verificación de firma dinámica

La verificación de la firma dinámica es una de las tecnologías biométricas conductuales más populares. En los últimos 40 años, la verificación de la firma dinámica ha evolucionado de forma muy significativa y ha demostrado ser uno de los sistemas biométricos más fiables y convenientes en muchos sectores relevantes. Como rasgo biométrico del comportamiento, hay muchos factores que afectan al rendimiento de la firma en línea, como la ergonomía, la calidad del dispositivo de adquisición, la interoperabilidad del dispositivo [22], factores de usabilidad [23], el uso del dedo como herramienta de escritura [24], el efecto del envejecimiento [25], la calidad de la firma [26], la complejidad de la firma [27], la limitada cantidad de bases de datos públicas que ha motivado la generación de firmas sintéticas, etc. La mayoría de estos factores se revisan en [28].

En este tipo de sistemas la firma proporcionada es una firma digital que incluye la información temporal del proceso de firma. De esta forma, no sólo obtenemos la imagen final que compondría nuestra firma sino parámetros que nos indican como se ha realizado este proceso. Entre los parámetros temporales que se pueden capturar encontramos:

- **Coordenadas X e Y**: Punto de la pantalla sobre el que se encuentra el stylus.
- **Presión**: La presión ejercida en cada momento, no disponible si la firma se ha realizado con el dedo.
- **Ángulo de inclinación**: Ángulo con el que el stylus incide sobre la pantalla. Las firmas realizadas con el dedo no disponen de esta información.

- **Información de vuelo**: Información acerca de si el stylus está o no en contacto con la pantalla y en caso de no estarlo, coordenadas del mismo. En las firmas realizadas con el dedo sólo es posible capturar si hay contacto o no con la pantalla.

Los sistemas de verificación dinámicos se centran en el uso de estas series temporales para realizar el proceso de verificación, haciendo uso de esta forma de la mayor cantidad de información recopilada posible.

2.3.1. Time-Aligned Recurrent Neural Network

Introducción

En esta sección se explicará en detalle la arquitectura Time-Aligned Recurrent Neural Network (TA-RNN) propuesta en el Trabajo de Fin de Grado [1] y cuya eficacia ha sido demostrada en varios trabajos [29, 30].

En concreto, se explicará el sistema desarrollado durante el Trabajo de Fin de Grado debido a que es el punto de partida en el que se basa el desarrollo de sistemas de verificación de firma dinámica realizado durante este trabajo. Además, este sistema se usará como baseline para comparar los resultados obtenidos.

La arquitectura TA-RNN se diseñó con la idea de maximizar la diferencia entre las características de una firma genuina y sus falsificaciones, sin alterar su similitud con el resto de firmas genuinas pertenecientes al mismo usuario. De esta forma, al conseguir unas características más diferenciadas se facilita el proceso de comparación, mejorando de esta forma los resultados obtenidos.

Esta arquitectura cuenta con dos etapas, un primera etapa de alineamiento temporal de las características mediante DTW y una segunda etapa conformada por una red neuronal, que realiza el proceso de comparación, devolviendo así el resultado de la verificación. Las características de entrada al sistema son un conjunto de 23 funciones temporales por firma. Dichas funciones temporales se extraen a partir de las coordenadas X e Y y la presión. En el capítulo 4 se hablará en detalle de ellas.

Preprocesado de las funciones temporales

Antes de introducir las funciones temporales de las firmas a la red neuronal se realiza un preprocesado previo de las mismas. Este preprocesado consiste en un alineamiento de las funciones temporales mediante DTW. El alineamiento sólo se realiza para la firma de test, la cuál es alineada teniendo como referencia la firma genuina.

De esta manera, si la firma de prueba no pertenece al mismo usuario que la genuina sus funciones temporales se distorsionarán alargándose o acortándose, haciendo que la diferencia entre ambas firmas sea más clara. En el caso de que la firma perteneciese al mismo usuario que la genuina, este alineamiento apenas realizaría variaciones, por lo que no cambiaría el parecido entre ambas.

Una vez alineadas las funciones temporales estas se envían a la primera capa de la red neuronal.

Red neuronal de comparación

La segunda etapa de este sistema está compuesto por una una red neuronal siamesa que tiene como entrada las funciones temporales alineadas de las firmas y obtiene a su salida una

puntuación entre 0 (genuina) y 1 (impostora) que expresa el resultado de la verificación. Esta red está a su vez compuesta por 3 capas, las cuáles podemos observar en la figura 4.3, que nos muestra la arquitectura de la red neuronal.

- La primera capa de la red está formada dos capas recurrentes ocultas, con 46 bloques de memoria cada una. Estas dos capas se encuentran en configuración siamesa, es decir, ambas capas comparten los mismos pesos, por lo que a efectos prácticos son la misma.
- La segunda capa es un capa recurrente oculta que tiene como entrada la salida concatenada de las dos capas anteriores y cuenta con 23 bloques de memoria.
- Finalmente, la tercera capa es una capa densa tipo feed-forward con activación sigmoidea que se encarga de extraer la puntuación final.

Protocolo Experimental

Para el entrenamiento de este sistema se usó la base de datos DeepSignDB, la misma base de datos que se utiliza en este trabajo y que se explicará en el capítulo 3. Esta base de datos se dividió a su vez en dos datasets distintos, uno para el desarrollo del sistema y otro para la evaluación del mismo. Esta división se hizo a nivel de usuario, no de firma, dejando el 70 % de los usuarios en el set de desarrollo y el 30 % en el set de evaluación.

A su vez, el set de desarrollo se dividió en dos subsets diferentes, uno de validación (20 %) y otro de entrenamiento (80 %), dejando de esta forma 867 usuarios para el entrenamiento.

Esta sistema fue entrenado de tres maneras diferentes atendiendo al tipo de falsificaciones utilizadas. En este trabajo sólo se utiliza como sistema baseline el que mejor resultados obtuvo. Este método consistió en utilizar el mismo número de *skilled forgeries* que de *random forgeries* para el entrenamiento del modelo. De esta manera, se obtiene un modelo robusto frente a ambos tipos de falsificaciones. El entrenamiento del modelo se hizo únicamente con firmas realizas con stylus, a pesar de que luego se comprobó su rendimiento con firmas realizadas con el dedo.

En la evaluación del sistema se utilizaron las otras 442 firmas restantes. Se evaluó el rendimiento del sistema sobre los dos tipos de falsificaciones (*skilled y random forgeries*) y en dos casos distintos, comparaciones 1vs1 y comparaciones 4vs1. Las comparaciones 1vs1 tienen en cuenta una única comparación genuina-impostora, sin embargo, las comparaciones 4vs1 promedian el resultado de la comparación de una firma de test con cuatro firmas genuinas del mismo usuario. Estas pruebas se realizaron sobre las firmas realizadas con stylus y las firmas realizadas con el dedo.

Resultados

La tabla 2.3.1 muestra una comparativa de diferentes sistemas de verificación de firma dinámica basados en deep learning. Se muestra el rendimiento de cada sistema en términos de EER para un tipo de falsificación concreto. También se indica la base de datos usada en cada caso, el número de firmas de entrenamiento y el tipo de sistema. A excepción de algunos sistemas, para la que se muestran los resultados sobre firmas realizadas con el dedo, sólo se muestran resultados obtenidos sobre firmas realizadas con stylus.

Como podemos observar TA-RNN obtiene unos resultados muy notables que prácticamente superan al del resto de propuestas. Estos resultados se tomarán como referencia a lo largo de este trabajo y se compararán con los resultados que obtengan los sistemas de verificación dinámica desarrollados.

Study	**Classifiers**	**Database**		**Experimental Protocol**			**Performance (EER)**
		Name	**# Users**	**# Train Users**	**Input**	**# Train Sig.**	
Otte *et al.* (2014) [31]	LSTM	SigComp2011	20	20	Stylus	12	Skilled = 23.8 %
Tolosana *et al.* (2018) [32]	BLSTM/BGRU	BiosecurID	400	300	Stylus	1	Skilled = 6.8 % Random = 5.4 %
						4	Skilled = 5.5 % Random. = 2.9 %
Lai and Jin (2018) [33]	GARU + DTW	MCYT	100	80	Stylus	5	Skilled = 1.8 % Random = 0.2 %
		Mobisig	83	70	Finger	5	Skilled = 10.9 % Random = 0.6 %
		e-BioSign	65	30	Stylus	4	Skilled = 6.9 % Random = 0.4 %
Ahrabian and Babaali (2018) [34]	LSTM Autoencoder	SigWiComp2013	31	11	Stylus	5	Skilled = 8.7 % Random = Unknown
Hefny and Moustafa (2019) [35]	MLP	SigComp2011	64	-	Stylus	5	Skilled = 0.5 % Random = Unknown
Wu *et al.* (2019) [36]	CNNs + DTW	MCYT	100	50	Stylus	5	Skilled = 2.4 % Random = Unknown
Li *et al.* (2019) [37]	LSTM	BiosecurID	132	110	Stylus	1	Skilled = 3.7 % Random = 1.9 %
		MCYT	100	85	Stylus	1	Skilled = 10.5 % Random = Unknown
		SCUT-MMSIG	50	40	Stylus	1	Skilled = 13.9 % Random = Unknown
		Mobisig	83	70	Finger	1	Skilled = 16.1 % Random = Unknown
Sekhar *et al.* (2019) [38]	CNNs	MCYT	100	95	Stylus	1	Skilled = 93.9 % Acc. Random = Unknown
		SVC-Task 2	40	35	Stylus	1	Skilled = 77.0 % Acc. Random = Unknown
Lai *et al.* (2020) [39]	CNNs	MCYT	100	90	Stylus	5	Skilled = 1.7 % Random = Unknown
		SVC-Task 2	40	36	Stylus	5	Skilled = 4.6 % Random = Unknown
Nathwani (2020) [40]	BLSTM/BGRU	SVC	-	-	Stylus	-	Skilled = 8.8 % AE Random = Unknown
Proposed	**TA-RNNs**	**DeepSignDB**	**1526**	**1084**	**Stylus**	**1**	**Skilled = 4.2 %** **Random = 1.5 %**
						4	**Skilled = 3.3 %** **Random = 0.6 %**
					Finger	**1**	**Skilled = 13.8 %** **Random = 1.8 %**
						4	**Skilled = 11.3 %** **Random = 1.0 %**

Cuadro 2.1: Comparativa del sistema TA-RNN con diferentes aproximaciones basadas en deep learning para verificación de firma dinámica. Fuente [30].

2.4. Sistemas de verificación de firma estática

Los sistemas de verificación de firma estática utilizan imágenes de las firmas como dato de entrada. Al solo poder usar la imagen como entrada este tipo de sistemas tiene mucha menos información disponible sobre el proceso de firma, lo que dificulta la tarea de verificación y explica que los resultados en este campo sean peores que en el de las firmas dinámicas.

Principalmente, las bases de datos utilizadas para entrenar estos sistemas contienen firmas manuscritas hechas en papel, y es por ello que solo se dispone de la imagen Sin embargo, también existe otro tipo de firmas estáticas que pueden ser utilizadas por estos sistemas. Se trata de las firmas sintéticas. Estas firmas se obtienen a partir de las firmas dinámicas, procesando las señales temporales y obteniendo la imagen final que resultaría. De esta manera podemos ampliar en gran medida los datos de entrenamiento de estos sistemas.

2.4.1. Extracción de características para verificación de firmas a través de Redes Neuronales Convolucionales

Introducción

En esta sección se describirá un sistema de verificación de firma estática del estado del arte propuesto por Luis G. Hafemann, Robert Sabourin y Luiz S. Oliveira en [41]. De este trabajo se usarán algunas partes para el desarrollo de sistemas de verificación de firma estática, en concreto, la fase de extracción de características. Sin embargo, se explicará el sistema completo con el fin de aportar todo el contexto.

El sistema está compuesto por dos partes, un extractor de características y un comparador de firmas. El extractor de características a su vez está formado por una red neuronal convolucional (*en inglés CNN*), un tipo de red que obtiene buenos resultados en problemas relacionados con imágenes. Por otro lado, para la comparación de las firmas y la obtención del resultado de la verificación se utiliza un SVM específico para cada usuario. Por lo tanto, este sistema requiere del entrenamiento de un SVM para cada usuario registrado.

La idea subyacente a dicho esquema es tratar de crear un espacio de características robusto a partir del cuál poder entrenar clasificadores para cada usuario. De esta forma, si se extraen las características más diferenciadoras de las firmas, la comparación entre estas se facilita.

Con el fin de emular la realidad, el grupo de usuarios empleado para el entrenamiento de la red de extracción de características es distinto al utilizado para entrenar los clasificadores. Gracias a ello, verificamos que el extractor de características generaliza correctamente frente a nuevos usuarios.

El sistema está formado por tres etapas: preprocesado, extracción de características y entrenamiento de los clasificadores. A continuación, se describirán dichas etapas.

Preprocesado

El conjunto de datos usado para el desarrollo de este sistema son firmas manuscritas que se han extraído de los documentos donde estaban escritas. Al no tener estas firmas un formato común y al tener la red una entrada estándar, es necesario normalizar las firmas. De esta manera conseguimos que cumplan con el formato de entrada y a la vez tengan el mismo rango de valores.

Inicialmente se centran las imágenes usando su centro de masa. Después, mediante el algoritmo OTSU se elimina el fondo dejando los píxeles de fondo en blanco y los píxeles de la firma en escala de grises. Finalmente, se invierte la imagen restándola a la intensidad máxima (255), y se redimensiona al tamaño de entrada de la red.

Red de extracción de características

Esta sección describe la red de extracción de características. En la figura 2.2 apreciamos su arquitectura. Vemos que no sólo cuenta con las capas de extracción de características propiamente dichas sino que además cuenta con dos capas de clasificación. Estas capas son las que fuerzan a la red a aprender características diferenciadoras, de forma que las firmas sean linealmente separables en el nuevo espacio de características creado.

La entrada de la red es una imagen de dimensiones $1 \times 150 \times 220$, es decir, una imagen en escala de grises. Después encontramos cinco capas convolucionales intercaladas con capas de max-pooling, reduciendo el tamaño de los datos hasta llegar a las capas fully-connected. Estas

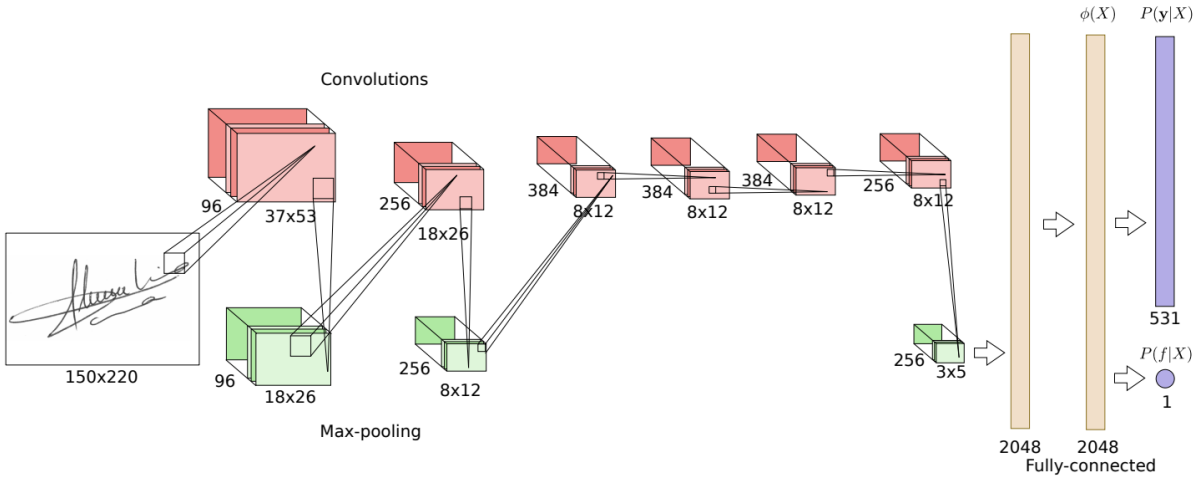


Figura 2.2: Ilustración de la arquitectura de la red de extracción de características. La imagen de entrada pasa por una secuencia de transformaciones con capas convolucionales, capas de max pooling y capas fully connected. La CNN se utiliza para proyectar las imágenes de la firma en otro espacio de características (análogo a extraer características), realizando una propagación feed-forward hasta una de las últimas capas antes de la capa de clasificación final, obteniendo el vector de características $\phi(X)$. Fuente [41].

capas tienen 2048 neuronas y a su salida proporcionan un vector de 1×2048, de esta manera hemos conseguido transformar 33000 píxeles en un vector de 2048 elementos.

Finalmente, el vector de características se envía a una capa con activación softmax la cuál clasifica la firma en uno de los usuarios de entrenamiento. Adicionalmente existe otra capa que permite discernir en último lugar si la firma es o no una falsificación. Se trata de una capa fully-connected con activación sigmoidea.

Gracias a la existencia de la última capa de activación sigmoidea tenemos dos formas de entrenar la red dando lugar a dos modelos diferentes:

- **Modelo Signet**: Este modelo resulta del entrenamiento del sistema únicamente con firmas genuinas, no usando la última capa sigmoide.
- **Modelo Signet-F**: Obtenemos este modelo al entrenar el sistema usando también las falsificaciones de las firmas, haciendo uso por lo tanto de la última capa.

Clasificadores dependientes del usuario

Tras el entrenamiento de la red de extracción de características se obtiene un vector de 2048 elementos para cada usuario. Con este vector se entrena un modelo SVM específico para cada usuario del conjunto designado para este entrenamiento, con usuarios distintos a los empleados en el entrenamiento de la red.

Cada modelo SVM es entrenado con un conjunto de firmas genuinas del usuario, que conformarán las muestras positivas, y un conjunto de firmas genuinas de otros usuarios, las muestras

negativas. Para el entrenamiento de estos modelos se puede usar tanto una función de base radial (RBF Kernel) como un función lineal (Linear SVM).

Protocolo experimental

Para realizar los experimentos se utilizaron las bases de datos GPDS-960 [42], MCYT-75 [43], CEDAR [44]. y PUC-PR brasileño [45]. Principalmente se utilizó GPDS-960 para el entrenamiento, debido a que es el mayor conjunto de datos disponible para la verificación de firma estática, contando con 881 usuarios.

Esta base de datos se separó en dos conjuntos, desarrollo y explotación. Se utiliza un subconjunto de usuarios de esta base de datos para aprender las características (el conjunto de desarrollo D) y otro para evaluar cómo estas características se generalizan a otros usuarios en este conjunto de datos (el conjunto de explotación E). El resto de bases de datos (MCYT, CEDAR y PUC-PR) se utilizan sólo para evaluar si las características generalizan a otros conjuntos de datos, por lo que sólo se usan en el mismo contexto que el conjunto de datos de explotación.

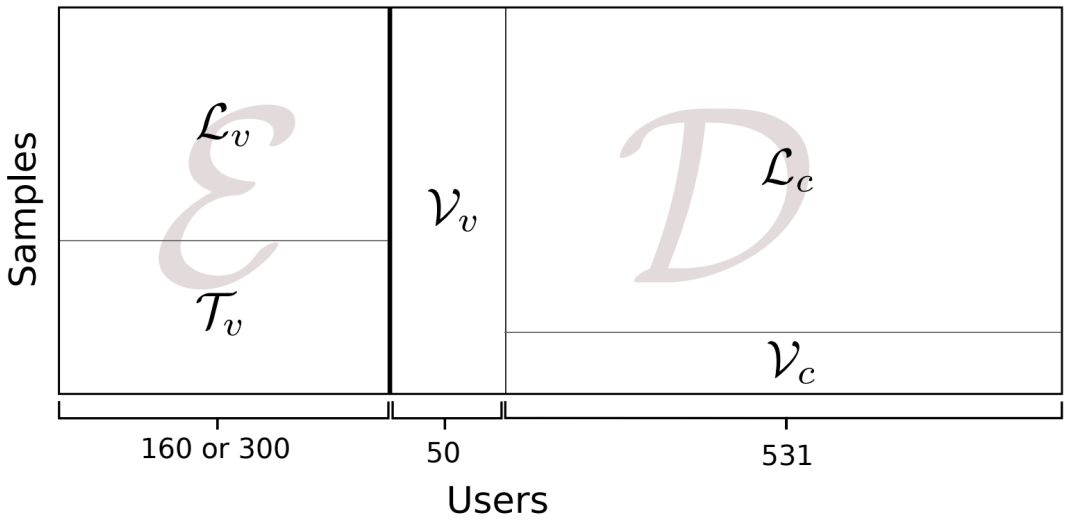


Figura 2.3: La base de datos GPDS se separa en conjunto de explotación ξ y conjunto de desarrollo D. El conjunto de desarrollo se usa para el entrenamiento de la red de extracción de características. El conjunto de explotación representa a los usuarios registrados en el sistema, estos usuarios se utilizan para entrenar los clasificadores SVM. Fuente [41].

El conjunto de datos de GPDS se divide de la siguiente manera, como se ilustra en la Figura 2.3. El conjunto de desarrollo utilizado para entrenar las redes convolucionales se divide en entrenamiento (90 %) y validación (10 %).

Una vez entrenadas las CNN, se entrenan los clasificadores SVM en un conjunto de validación *Vv* (conjunto de validación para la verificación) compuesto por 50 usuarios. Dicho se utiliza para estimar el rendimiento de los clasificadores entrenados con el espacio de características aprendido por la CNN. Se usan 12 firmas genuinas de un usuario como muestras positivas y 12 firmas de otros usuarios de *Lc* como muestras negativas.

Por último, se utilizan los modelos e hiperparámetros que mejor funcionaron en el conjunto de validación para entrenar y probar los clasificadores sobre el conjunto de explotación ξ. Se entrenan los SVM con el conjunto *Lv* y se evaluan con el conjunto *Tv*. Para cada usuario, se usa

un conjunto de datos que consiste en r firmas genuinas del usuario como muestras positivas, y firmas genuinas de otros usuarios como muestras negativas. Se han probado distintos números de firmas de entrenamiento.

La evaluación del rendimiento en el conjunto de pruebas se ha realizado usando las siguientes métricas:

- **Tasa de falsos rechazos (FRR)**: la fracción de firmas genuinas rechazadas como falsas.
- **Tasa de falsos positivos (FARrandom y FARskilled)**: la fracción de falsificaciones aceptadas como auténticas (considerando las falsificaciones aleatorias y las falsificaciones intencionadas).
- **Equal Error Rate (EER)**: que es el error obtenido cuando FAR = FRR.

Resultados

A continuación, se muestran los resultados obtenidos por este sistema sobre la base de datos GPDS. Los resultados han sido obtenidos usando firmas estáticas escritas sobre papel, por lo tanto no sintéticas. En la tabla 2.2 podemos apreciar los resultados de este sistema comparados con otros del estado del arte. Estos resultados están expresados en términos de EER y sobre *skilled forgeries*.

Sistema	Dataset	#muestras por usuario	Características extraídas	EER (%)
Hu and Chen [46]	GPDS-150	10	LBP, GLCM, HOG	7.66
Guerbai *et al.* [47]	GPDS-160	12	Curvelet transform	15.07
Serdouk *et al.* [48]	GPDS-100	16	GLBP, LRF	12.52
Yilmaz [49]	GPDS-160	5	LBP, HOG, SIFT	7.98
Yilmaz [49]	GPDS-160	12	LBP, HOG, SIFT	6.97
Soleimani *et al.* [50]	GPDS-300	10	LBP	20.94
Hafemann *et al.* [41]	GPDS-160	5	SigNet	3.23 (+-0.36)
Hafemann *et al.* [41]	GPDS-160	12	SigNet	2.63 (+-0.36)
Hafemann *et al.* [41]	GPDS-300	5	SigNet	3.92 (+-0.18)
Hafemann *et al.* [41]	GPDS-300	12	SigNet	3.15 (+-0.18)
Hafemann *et al.* [41]	GPDS-160	5	SigNet-F	2.41 (+-0.12)
Hafemann *et al.* [41]	GPDS-160	12	SigNet-F	2.41 (+-0.12)
Hafemann *et al.* [41]	GPDS-300	5	SigNet-F	2.42 (+-0.24)
Hafemann *et al.* [41]	GPDS-300	12	SigNet-F	1.69 (+-0.18)

Cuadro 2.2: Comparación con sistemas del estado del arte sobre la base de datos GPDS. Fuente [41].

Vemos que el sistema mejora notablemente el rendimiento del resto de modelos, rebajando el EER hasta el 1.72 ($\pm$ 0.15) % utilizando 12 firmas por usuario. Debido a estos resultados, se utilizará la red de extracción de características de este sistema para el desarrollo de sistemas de verificación de firma estática, con la diferencia de que sólo se usarán firmas sintéticas.

2.5. Sistemas de detección de la complejidad

La complejidad es un factor importante en algunos sistemas de autenticación tradicionales, como los basados en contraseñas, en los que se necesita una complejidad mínima de la contraseña para garantizar un nivel mínimo de seguridad. Los sistemas de verificación de firma dinámica también pueden sufrir este problema. El análisis de la complejidad de las firmas permite averiguar cómo afecta la complejidad de las firmas al rendimiento de un sistema de verificación. Además, sería posible advertir a los usuarios con firmas vulnerables en cuanto a su complejidad, para que pudieran modificar su firma por otra más robusta.

El efecto de la complejidad de la firma en el rendimiento del sistema ha demostrado ser un factor importante. En [51], Alonso-Fernández *et al.* se evaluó el efecto de la complejidad y legibilidad de las firmas para la verificación de firmas off-line (es decir, firmas sin información dinámica disponible) señalando las diferencias de rendimiento para varios matchers.

Se han propuesto diferentes enfoques para detectar la complejidad de las firmas manuscritas dinámicas. Houmani *et al.* propuso una medida de entropía basada en la estimación de la densidad local mediante un modelo de Markov oculto (HMM) [52]. Miguel-Hurtado *et al.* investigó la creación de un novedoso modelo matemático para la evaluación automática de la complejidad de la firma [53]. Una aproximación reciente fue realizada por Tolosona *et al.* en [27]. Este estudio propuso un detector de complejidad basado en el uso del número de trazos aplicando el conocido modelo Sigma LogNormal de generación de escritura [54]. Posteriormente, una vez clasificadas las firmas en tres niveles de complejidad (baja, media y alta), se extrajeron las funciones de tiempo óptimas asociadas a cada nivel de complejidad específico. Finalmente, midieron el rendimiento de un sistema de verificación de firmas basado en Dynamic Time Warping (DTW) utilizando únicamente las funciones de tiempo óptimas para cada clase de complejidad. De este modo, se obtuvieron mejoras significativas en el rendimiento del sistema en comparación con un sistema basado en DTW que utiliza las mismas funciones temporales para todas las firmas sin tener en cuenta su nivel de complejidad.

Recientemente, Vera-Rodríguez *et al.* han propuesto en [55] un nuevo sistema de detección de complejidad basado en técnicas de Deep Learning (DL). Este sistema se desarrolló a través de un proceso semisupervisado en el que se entrenó un modelo inicial sobre una base de datos de tamaño medio (BioSecurID [56]), que luego se utilizó para clasificar la complejidad de las firmas de una base de datos mucho mayor (DeepSignDB [19]). Finalmente, a partir de estas etiquetas automáticas desarrollaron el sistema de detección de complejidad, logrando resultados de alrededor del 85 % de precisión en comparación con las etiquetas manuales. Este sistema mejoró significativamente los resultados del 64 % de precisión conseguidos anteriormente en [27] en las mismas condiciones experimentales.

En concreto, este último sistema será empleado en la primera parte de este proyecto para la detección de la complejidad de las firmas, permitiendo así el análisis de los efectos que dicha complejidad ocasiona en sistemas de verificación de firma dinámica.

2.6. Fusión de sistemas dinámicos y estáticos

Como hemos visto, los sistemas de firma dinámica y los sistemas de firma estática hacen uso de características diferentes de las firmas para llevar a cabo la verificación de las mismas. Esto nos hace preguntarnos si estas características extraídas son independientes y la combinación de la información obtenida por ambos tipos de sistemas se puede fusionar obteniendo así un mejor rendimiento.

El primer problema que se nos plantea es la necesidad de tener una copia de cada firma en versión dinámica y estática. Este problema es uno de los principales factores que dificultan la combinación de sistemas. Sin embargo, la solución a este problema viene dada por la creación de firmas sintéticas a partir de los datos dinámicos de las firmas, de manera que podemos tener por cada muestra dinámica una muestra estática sintética. Una vez creada nuestra base de datos sintética, se nos plantea la elección del método de fusión. Las dos opciones posibles son la combinación de las puntuaciones resultantes de cada sistema y la combinación de las características de ambos sistemas.

Un trabajo que resulta interesante analizar cuando hablamos de combinación de sistemas es el que llevaron a cabo Galbally *et al.* en [57]. En este trabajo los autores realizaron una comparación de rendimiento de los principales sistemas dinámicos y estáticos del estado del arte para posteriormente analizar la complementariedad de la firma dinámica y estática en los escenarios de falsificación *random* y *skilled*. Además, desarrollaron un nuevo método para la generación de firmas sintéticas y propusieron una arquitectura para verificación de firma dinámica basada en la combinación de información dinámica y estática sintética.

Gracias a su nuevo método de generación de firmas sintéticas solventaron la necesidad de tener por cada firma una muestra dinámica y otra estática. Mediante la realización de sus experimentos concluyeron que el mayor aporte que la combinación de sistemas aportaba para el rendimiento de sistemas de verificación dinámicos se obtenía sobre las falsificaciones *random*. Para los tres sistemas que probaron obtuvieron un notable aumento del rendimiento sobre este tipo de falsificaciones, de un 40 % en media. Por otro lado, en la detección de falsificaciones *skilled* la combinación de sistemas no mejoraba los rendimientos, simplemente estos se mantenían o disminuían muy levemente. Únicamente en el caso de que el sistema dinámico combinado tuviera previamente un bajo rendimiento sobre falsificaciones *skilled* se lograba una mejora debida a la fusión.

Esta investigación refuerza las conclusiones de trabajos anteriores que muestran que, aunque la firma dinámica tiene un mayor potencial para tareas de verificación, no comprende toda la información presente en el rasgo de la firma. De este modo, la información estática puede ser un activo muy valioso para aumentar significativamente el rendimiento general de este rasgo biométrico.

3

Bases de Datos

3.1. Introducción

En este apartado se detallarán las bases de datos utilizadas a lo largo del proyecto. Para el desarrollo de los sistemas de firma dinámica se ha usado la base de datos DeepSign, tanto las firmas hechas con stylus como las firmas realizadas con el dedo. Para el desarrollo de sistemas de firma estática se ha utilizado esta misma base de datos pero con la versión sintética estática de las mismas. Esta versión sintética se obtiene a partir de las series temporales de las firmas dinámicas.

A continuación, se describirá brevemente el contenido de cada una de las bases de datos utilizadas.

3.2. Base de datos de firma dinámica

Como hemos comentado antes, la base de datos utilizada para el entrenamiento y prueba de los sistemas de firma dinámica desarrollados en este trabajo ha sido DeepSignDB [19]. DeepSign es una base de datos muy amplia que cuenta con 1526 usuarios, lo que supone un total de más de 50000 firmas. Entre las firmas que integran DeepSign encontramos tanto firmas realizadas con stylus como firmas realizadas con el dedo. El gran volumen de usuarios con el que cuenta esta bases de datos es debido a que está compuesta a su vez por cinco bases de datos diferentes, cuya información aparece descrita en la tabla 3.1. Ahora explicaremos brevemente cada una de estas cinco bases de datos:

- **MCYT** [58]: La base de datos MCYT está compuesta por un total de 25 firmas genuinas y 25 falsificaciones por usuario, adquiridas en una única sesión en bloques de 5 firmas. Hay un total de 330 usuarios cuyas firmas fueron adquiridas en un entorno controlado. El dispositivo de captura usado fue la tablet Wacom Intuos A6 que capturó, con una resolución de 100Hz, las coordenadas espaciales X e Y (con una resolución de 0.25mm), la presión (1024 niveles) y la orientación angular del bolígrafo. Además también está disponible la información referente a los pen-ups.

En cuanto al tipo de impostores, se consideraron solo falsificaciones estáticas, permitiendo a los falsificadores tener acceso a la imagen de la firma a falsificar.

- **BiosecurID** [56]: Esta base se compone de un total de 16 firmas genuinas y 12 falsificaciones por usuario, capturadas en 4 sesiones distintas separadas por un intervalo de 2 meses entre ellas. Hay un total de 400 usuarios cuyas firmas fueron adquiridas en un entorno controlado. El dispositivo de captura usado fue la tableta Wacom Intuos 3 que capturó, con una resolución de 100Hz, las coordenadas espaciales X e Y (con una resolución de 0.25mm), la presión (1024 niveles) y la orientación angular del bolígrafo. En esta base de datos también está disponible la información referente a los pen-ups.

  En este caso se consideraron falsificaciones estáticas y dinámicas, permitiendo a los falsificadores tener acceso a la imagen de la firma a falsificar en las dos primeras sesiones y de las dinámicas de la firma en las dos últimas sesiones.

- **Biosecure DS2** [59]: La base de datos BioSecure DS2 está compuesta por un total de 30 firmas genuinas y 20 falsificaciones por usuario, capturadas en 2 sesiones dejando un periodo de tres meses entre las mismas. En total cuenta con 650 usuarios cuyas firmas fueron adquiridas en un entorno controlado. Se usó el mismo dispositivo de captura (Wacom Intuos 3) y las mismas condiciones de captura que en la recopilación de los datos de la base BioSecurID.

  En esta base de datos solo se consideraron falsificaciones dinámicas. Cabe destacar la alta calidad de las falsificaciones debido a que en este caso los falsificadores tuvieron acceso a la imagen y las dinámicas de las firmas.

- **e-BioSign DS1** [24]: La base de datos e-BioSign DS1 está compuesta por firmas capturadas con 5 dispositivos distintos. Tres de ellos están específicamente diseñados para capturar escritura (Wacom STU-500, STU-530, y DTU-1031), mientras que los otros dos son tablets de propósito general que no están diseñadas para esta tarea específica (Samsung ATIV 7 and Galaxy Note 10.1). Es importante remarcar que en cada dispositivo se usó con su propio stylus. Además en esta base de datos también contamos con firmas realizadas con el dedo para los dispositivos Samsung.

  Las firmas fueron recopiladas en dos sesiones para un total de 65 usuarios, dejando un periodo de tres semanas entre cada sesión. Para cada usuario encontramos 8 firmas genuinas y 6 falsificaciones.

  Se usaron falsificaciones estáticas y dinámicas en ambas sesiones, permitiendo a los falsificadores ver la imagen y las dinámicas de las firmas.

- **e-BioSign DS2**: En esta base de datos se ha seguido un protocolo de captura similar al de e-BioSign DS1. En concreto para capturar escritura con stylus se ha usado la tablet Wacom STU-530. Cuenta con 8 firmas genuinas y 6 falsificaciones por usuario. Por otro lado para capturar escritura realizada con el dedo se han usado la tablet Samsung Note 10.1 y el smartphone Samsung S3. En total se capturaron firmas de 81 usuarios en dos sesiones distintas con un intervalo de tiempo entre ambas de 3 semanas.

  A diferencia de e-BioSign DS1, en esta base de datos solo se consideraron falsificaciones dinamicas, permitiendo a los falsificadores ver tanto la imagen como las dinámicas de las firmas para lograr falsificaciones de alta calidad.

STYLUS

Base de datos	Usuarios	Dispositivos de captura	Nº de Firmas Genuinas	Nº de Falsificaciones
MCYT	330	Wacom Intuos A6	25	25
BioSecurID	400	Wacom Intuos 3	16	12
BioSecure DS2	650	Wacom Intuos 3	30	20
e-BioSign DS1	65	Wacom STU-500 Wacom STU-530 Wacom DTU-1031 Samsung ATIV 7 Samsung Note 10.1	8	6
e-BioSign DS2	81	Wacom STU-530	8	6

DEDO

Base de datos	Usuarios	Dispositivos de captura	Nº de Firmas Genuinas	Nº de Falsificaciones
e-BioSign DS1	65	Samsung ATIV 7 Samsung Note 10.1	8	6
e-BioSign DS2	81	Samsung Note 10.1 Samsung S3	8	6

Cuadro 3.1: Características de las bases de datos recogidas en DeepSignDB. Los números de firmas son por usuario y dispositivo.

3.3. Base de datos de firma estática

Para el desarrollo de sistemas de verificación de firma estática se han utilizado firmas sintéticas obtenidas a partir de la base de datos DeepSign. Esta base de datos se compone de las series temporales que muestran la evolución de las coordenadas X e Y y de la presión de cada firma. Sólo se ha creado una versión sintética de las firmas realizadas con stylus, debido a que las firmas realizadas con el dedo no se han utilizado para en el desarrollo de estos sistemas.

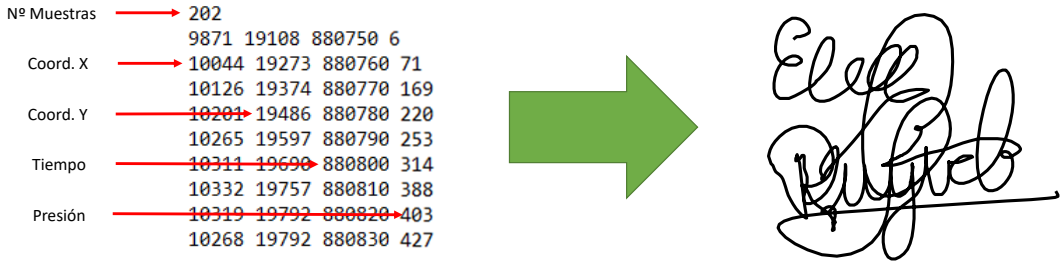


Figura 3.1: Muestras temporales de la firma dinámica (izquierda) y la firma estática resultante del proceso (derecha).

El proceso de creación de la base de datos de firmas estáticas sintéticas se ha realizado mediante un script en Matlab. En este script se dibuja la firma en función de las series temporales y atendiendo a tres parámetros:

- **Tamaño de píxel**: Con este parámetro se indica el grosor del trazo de la firma. En concreto, se ha utilizado un grosor de píxel 3.
- **Información de vuelo**: Indica si en la imagen final se tiene en cuenta el movimiento del stylus mientras no está en contacto con la pantalla, es decir, mientras no está dibujando. Se ha decidido no incluir esta información para asemejar las imagen de las firmas lo más posible a una firma real.
- **Presión**: Este parámetro nos permite variar el grosor del trazo en función de la presión ejercida, y de esta forma asemejar la imagen final a una firma real. En este trabajo no se ha utilizado esta funcionalidad.

Las decisiones tomadas a la hora de establecer los parámetros del script de obtención de firmas sintéticas han sido basadas en el trabajo previo [60]. En este trabajo se realizaron diversas pruebas variando los parámetros y entrenando después un sistema de verificación de firma estática compuesto por redes convolucionales. Los parámetros que obtuvieron mejor resultados son los utilizados en este proyecto.

4

Sistema propuesto

4.1. Introducción

El desarrollo de este trabajo se ha llevado a cabo tres etapas, cada una de las cuáles con un objetivo diferente, por lo tanto, con un sistema distinto. Sin embargo, todas las fases están alineadas en torno a un objetivo común: mejorar el rendimiento de un sistema de verificación de firma dinámica.

Inicialmente se llevó a cabo un análisis de los efectos que la complejidad de las firmas tenía en el rendimiento de sistemas de verificación de firma dinámica. Con la intención de mejorar el rendimiento de estos sistemas al introducir la información de complejidad. Después se buscó mejorar sistemas de verificación firma estática para finalmente combinar los sistemas estáticos y dinámicos y obtener un sistema combinado que obtuviese buenos resultados en la verificación de firma dinámica.

A continuación, se explican en detalle los sistemas resultantes de la labor desarrollada a lo largo de este trabajo para cada una de las tres fases.

4.2. Sistemas de análisis de complejidad

En esta sección se describe el enfoque propuesto para evaluar los efectos de la complejidad de las firmas en sistemas de verificación de firma dinámica. Con este propósito se ha diseñado una arquitectura basada en la detección de la complejidad de las firmas como paso previo a su paso por un sistema de verificación de firma dinámica basado en la complejidad, como se muestra en la Fig. 4.1.

La idea principal de este apartado es analizar cómo la complejidad de las firmas evaluadas afecta al rendimiento de un sistema de verificación y si es posible explotar esa información de complejidad para desarrollar un sistema de verificación de firma dinámica basado en la complejidad que pueda alcanzar mayor rendimiento. Para ello, se ha definido una arquitectura sencilla pero eficaz. En la verificación de firmas dinámica tendremos una firma de prueba (desconocida) que va a ser comparada con una firma genuina para verificar su identidad.

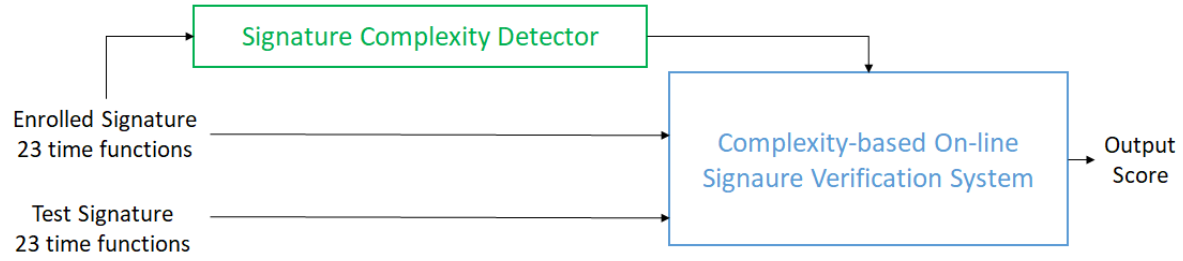


Figura 4.1: Arquitectura propuesta para un sistema de verificación de firma dinámica basado en complejidad.

En primer lugar, la firma inscrita pasa por un detector de complejidad de la firma que la clasifica en uno de los tres niveles de complejidad. A continuación, las dos firmas, junto con la etiqueta de complejidad de la firma registrada, entran en el sistema de verificación basado en la complejidad. Por último, el sistema de verificación proporciona un valor de puntuación de salida (entre 0 y 1), que indica la similitud de las dos firmas comparadas (se esperaría un valor cercano a 0 cuando las dos firmas son auténticas y un valor cercano a 1 cuando la firma de prueba es una falsificación hábil o aleatoria).

A continuación, describimos brevemente las dos tecnologías utilizadas tanto para la estimación de la complejidad de la firma como para la verificación de la firma dinámica, así como el tipo de datos de entrada a dichos sistemas.

4.2.1. Datos de entrada: funciones temporales

La información dinámica de las firmas que va a conformar la entrada de los dos sistemas está compuesta por 23 funciones temporales. Estas funciones se han obtenido a partir de las 3 funciones iniciales, coordenadas X e Y y la presión, con la que está compuesta la base de datos utilizada. El empleo de estas funciones viene dado por el buen rendimiento que han demostrado trabajos anteriores [32] al usar como datos de entrada un mayor número de funciones temporales. La extracción de estas funciones temporales se ha realizado a través de la herramienta Matlab. Cada una de estas funciones temporales viene descrita en la Tabla 4.1.

En el caso de las firmas realizadas con el dedo las funciones relacionadas con la presión ejercida se han puesto a cero, debido a que esta información no se encuentra disponible para este tipo de firmas.

#	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 utilizadas.

4.2.2. Detector de la complejidad

La primera etapa de nuestro sistema propuesto consiste en un detector de la complejidad de la firma. Para ello hacemos uso del sistema propuesto en [55]. Este sistema clasifica cada firma en uno de los tres niveles de complejidad considerados: bajo, medio y alto. En la Fig. 4.2 se muestran ejemplos de firmas en los tres niveles de complejidad considerados en este trabajo.

Este sistema se desarrolló mediante un proceso semi-supervisado. En primer lugar, se entrenó un modelo inicial sobre la base de datos BiosecurID [56], para la que se dispone de las etiquetas manuales de complejidad de las firmas como de baja, media y alta complejidad. A continuación, este modelo se utilizó para obtener automáticamente las etiquetas de complejidad de una base de datos mucho más grande con más de 1500 sujetos (DeepSignDB [19]). Finalmente, a partir de estas etiquetas automáticas, se desarrolló un nuevo sistema de detección de la complejidad de las firmas basado en RNNs.

Para cada firma, las señales relacionadas con las coordenadas espaciales X e Y y la presión se utilizan para extraer las 23 funciones temporales [32], que son la entrada del sistema. Estas funciones temporales se preprocesan siguiendo una normalización de media cero y desviación estándar unitaria. La red se compone de dos capas de memoria larga a corto plazo bidireccional (BLSTM) y una capa feed-forward con activación softmax, que proporciona una puntuación de salida para cada uno de los tres niveles de complejidad considerados. Tal y como se describe en [55], este sistema alcanza un 85 % de precisión en la clasificación del nivel de complejidad en comparación con la clasificación manual para un conjunto de datos de evaluación de la base de datos de BiosecurID [56].

4.2.3. Sistema de verificación de firma dinámica

El sistema de verificación de firmas basado en la complejidad está basado en el sistema TA-RNN (Time-Aligned Recurrent Neural Networks) propuesto en [29], [30]. Para la entrada del sistema, alimentamos la red con las 23 funciones temporales extraídas de la firma, que son las

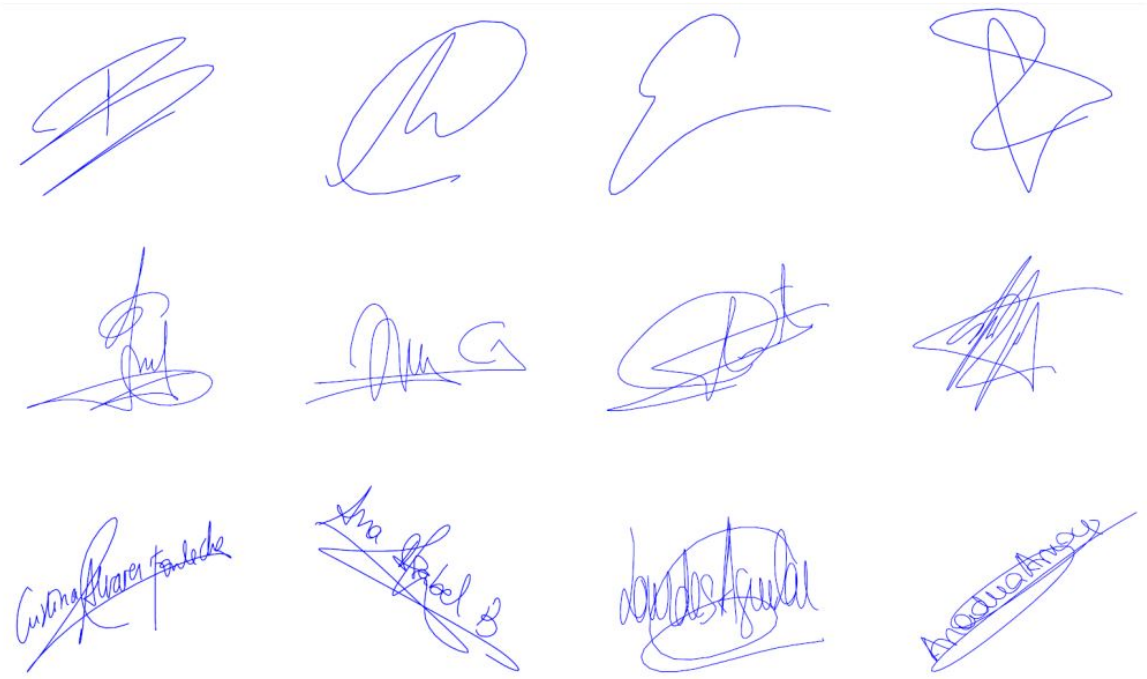


Figura 4.2: Ejemplos de firmas de los tres grupos de complejidad: baja complejidad (arriba), media complejidad (en medio) y alta complejidad (abajo).

mismas que se utilizan para el sistema de detección de complejidad. La arquitectura TA-RNN se basa en dos etapas: i) una primera etapa basada en la alineación de la secuencia temporal mediante DTW y ii) una segunda etapa consistente en una red neuronal, concretamente una RNN. El sistema desarrollado en [30] estaba compuesto por tres capas. La primera capa está compuesta por dos capas ocultas Bidirectional Gated Recurrent Unit (BGRU) con 46 bloques de memoria cada una, compartiendo los pesos entre ellas.

Las salidas de las dos primeras capas ocultas BGRU paralelas se concatena y sirve de entrada a la segunda capa, que corresponde a una capa oculta BGRU con 23 bloques de memoria. Por último, se considera una capa de red neuronal feed-forward con una activación sigmoidea que proporciona una puntuación de salida para cada par de firmas. En la Fig. 4.3 podemos observar esta arquitectura.

La arquitectura TA-RNN se utiliza en este trabajo como uno de los sistemas de referencia para comparar los resultados obtenidos con el sistema de verificación de firmas basado en la complejidad. En este sistema de verificación de firmas basado en la complejidad se llevan a cabo diferentes estrategias para explotar la complejidad de las firmas. En particular, se consideran tres enfoques principales: i) entrenar desde cero un modelo DL específico por nivel de complejidad, ii) entrenar un modelo DL específico por nivel de complejidad pero aplicando un ajuste fino a partir del modelo general de DeepSign en [30], y iii) entrenar sólo un modelo para todos los tipos de firmas, pero equilibrando las clases de complejidad durante el entrenamiento.

Los diferentes sistemas de verificación adaptados a cada complejidad utilizan la misma arquitectura que la descrita anteriormente. Sin embargo, para determinados experimentos se han realizado algunos cambios en la arquitectura. Estos detalles se describen en el capítulo siguiente.

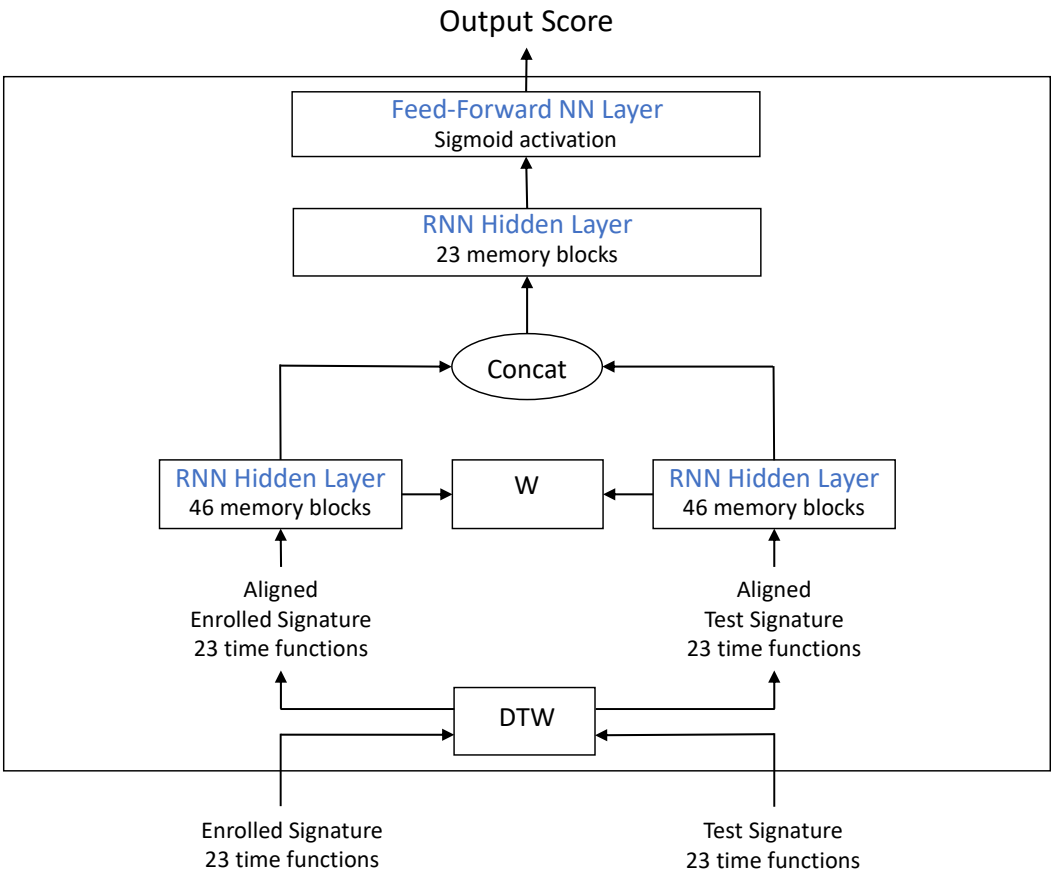


Figura 4.3: Sistema de verificación de firma dinámica.

4.3. Sistemas de verificación de firma estática

En esta sección se describen los sistemas resultantes de la búsqueda de mejora del rendimiento de los sistemas de verificación de firma estática. Para ello, se ha partido del sistema propuesto en [41] como sistema inicial. Este sistema obtenía buenos resultados sobre la base de datos de firma estática GPDS y sobre la base de datos de firma estática sintética MCYT. En esta ocasión se ha utilizado una base de datos con un mayor número de usuarios, DeepSignDB.

El sistema está formado por dos partes, una primera parte compuesta por una red neuronal con la cuál se realizaba la extracción de características de las firmas; y una segunda, basada en SVMs, en la que haciendo uso de esas características se entrenaba un clasificador que indicaba a qué usuario pertenecía una firma, o si no pertenecía a ningún usuario registrado.

Para este sistema no se ha usado el sistema propuesto en [41] al completo, en concreto, se ha hecho uso de la red convolucional que utilizaba para extraer las características de las firmas, a la que nos referiremos de ahora en adelante como red Signet. En la Fig. 4.4 podemos ver la parte del sistema utilizada. Una vez obtenido el rendimiento del sistema baseline, se han explorado dos arquitecturas principales, ambas integrando la CNN Signet, para intentar mejorar el rendimiento.

A continuación se describen las dos aproximaciones propuestas para el sistema de verificación de firma estática. Ambas arquitecturas hacen uso de la red Signet como extractora de

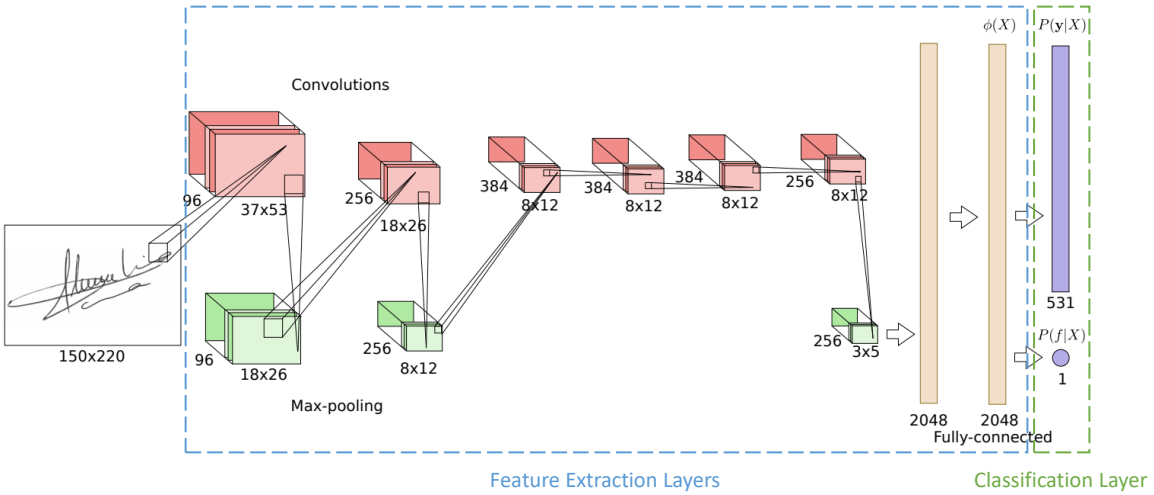


Figura 4.4: Ilustración de la arquitectura de la red de extracción de características. La CNN se utiliza para proyectar las imágenes de la firma en otro espacio de características (análogo a extraer características), realizando una propagación feed-forward hasta una de las últimas capas antes de la capa de clasificación final, obteniendo el vector de características $\phi(X)$. Fuente [41].

características.

4.3.1. Arquitectura Siamesa

Tras el buen resultado que los sistemas de arquitectura siamesa proporcionan para la verificación dinámica, se optó por esta arquitectura usando la red Signet como extractora de características, como podemos observar en la Fig. 4.5.

En esta arquitectura se calculan las características obtenidas por la red Signet para la firma genuina y la firma de test, una vez calculadas estas características se concatenan y se envían a una red de clasificación la cuál da como resultado un score entre 0 (genuina) y 1 (falsificación). Se han realizado experimentos con diferentes configuraciones para esta red, variando parámetros como el número de capas o el número de neuronas por capas. En el capítulo 5 se comentarán en detalle las distintas combinaciones. La configuración que mejor resultado arroja consta de la red Signet seguida de una capa de clasificación formada por 2 capas densa de 75 y 50 neuronas y una última capa densa con activación sigmoidea.

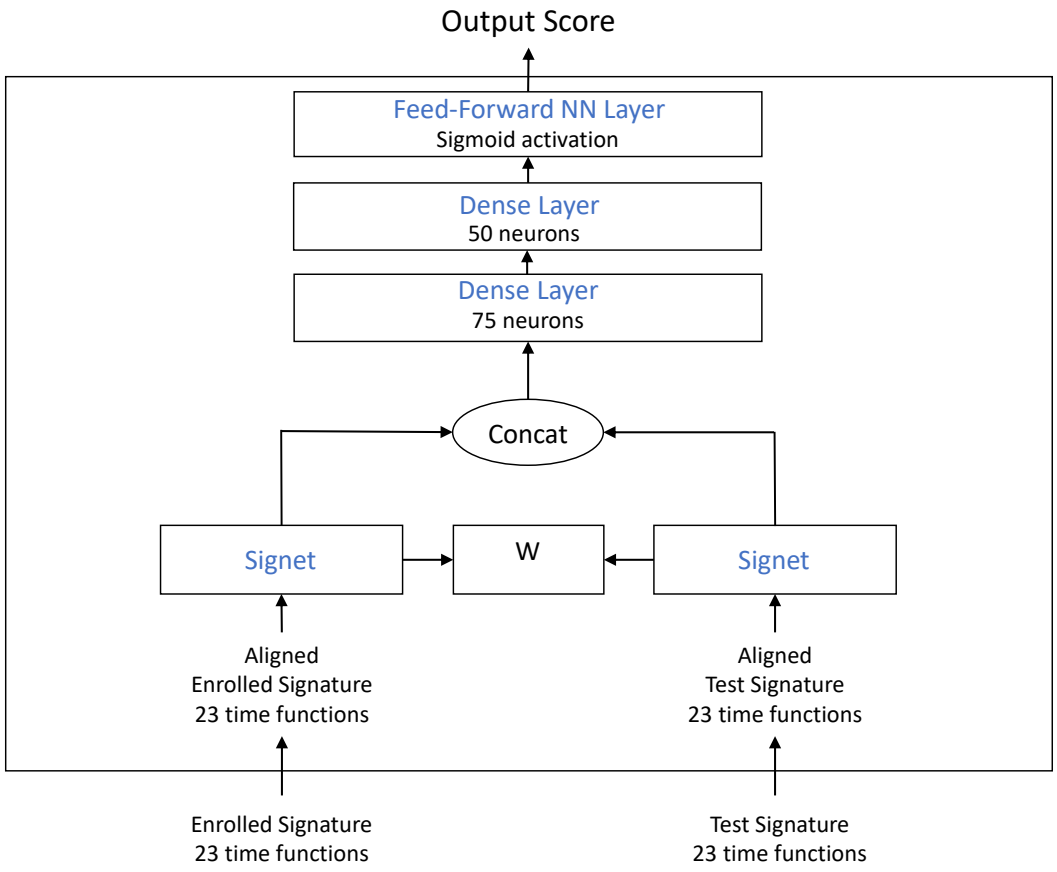


Figura 4.5: Sistema de verificación de firma estática con arquitectura siamesa.

4.3.2. Triplet Loss

En esta aproximación se utiliza el algoritmo triplet loss para, a partir de las características extraídas por la red Signet, obtener unas características más robustas con las que poder diferenciar mejor las firmas de cada usuario y facilitar de esta manera la comparación de ambas.

Triplet loss es una función de pérdidas que se utiliza en el entrenamiento de sistemas de machine learning para optimizar dicho entrenamiento. La función tiene como entrada tres tipos de muestras: una muestra que sirve como referencia y que recibe el nombre de ancla, una muestra positiva con la misma etiqueta que la muestra ancla, y finalmente, una muestra negativa con una etiqueta diferente a las otras dos. Para el caso de verificación, la muestra ancla y la positiva serían firmas genuinas del mismo usuario y la muestra negativa podría ser tanto una falsificación *skilled* de dicho usuario como una firma genuina de otro usuario (falsifiación *random*).

Esta función de pérdidas tiene como objetivo minimizar la distancia entre las muestras con la misma etiqueta y maximizar la distancia entre las muestras con etiquetas distintas. De esta forma, minimizaríamos la distancia entre las firmas genuinas de un mismo usuario y maximizaríamos la distancia entre la muestras genuinas de un usuario y sus falsificaciones, así cómo con el resto de firmas genuinas de otros usuarios. La función de pérdidas triplet loss viene definida por las siguientes ecuaciones:

$$L(A, P, N) = max(||f(A) - f(P)||^2 - ||f(A) - f(N)||^2 + \alpha, 0) \tag{4.1}$$

$$J = \sum_{i=1}^{M} L(A^{(i)}, P^{(i)}, N^{(i)}) \tag{4.2}$$

En ellas vemos como se minimiza la distancia entre las muestras ancla (A) y las muestras negativas (N), y se maximiza la distancia entre las muestra ancla y la muestras positivas (P). El factor α sirve como margen entre la distancia de las muestras ancla con las muestras positivas y las negativas.

Para el entrenamiento del sistema se parte de una red con tres entradas: dos firmas genuinas y una falsificación. De esta forma la red puede aprender a diferenciar mejor las falsificaciones, al tener una comparación genuina-genuina con la que compararse. Una vez acabado el entrenamiento los pesos de la red entrenada se guardan y se utilizan posteriormente en una arquitectura de verificación, que solo cuenta con dos firmas de entrada, compuesta por la capa de extracción de características entrenada con triplet loss y una capa de clasificación.

En este caso se han probado dos capas de clasificación diferentes: la distancia euclídea entre las características de las firmas de entrada y la capa de clasificación utilizada en la arquitectura siamesa.

4.4. Combinación de sistemas

Una vez desarrollados los sistemas de firma dinámica y estática, se ha procedido a la fusión de dichos sistemas mediante la combinación de las puntuaciones obtenidas por cada sistema, obteniendo así una puntuación global.

La combinación se ha realizado obteniendo la media entre las puntuaciones de los sistemas dinámico y estático. En el caso del sistema de firma estática con arquitectura siamesa la puntuación de salida es un número entre 0 (genuina) y 1 (falsificación), al igual que el sistema de firma dinámica. De esta forma, al tener dos sistemas con el mismo rango de salida se pueden combinar linealmente sus puntuaciones.

En cuanto a las distancias euclídeas, se ha procedido a normalizarlas utilizando el método min-max. Este método de normalización devuelve datos con un rango de salida 0 - 1, de forma que nos permite la combinación directa con las puntuaciones del sistema dinámico. Esta combinación también se ha realizado obteniendo la media entre ambas puntuaciones.

5

Resultados del análisis y explotación de la complejidad

5.1. Introducción

En este primer conjunto de experimentos se busca realizar un análisis en profundidad sobre cómo afecta a los sistemas de verificación de firma dinámica la complejidad de las firmas de entrada y si es posible usar esta información para desarrollar sistemas más eficientes.

En primer lugar, realizamos un análisis del rendimiento del sistema de verificación de firma dinámica utilizando la base de datos pública DeepSignDB para estudiar la diversidad de los sujetos en cuanto a la complejidad de sus firmas. A continuación, llevamos a cabo varios experimentos para explotar la complejidad de las firmas con el objetivo de mejorar el rendimiento general de un sistema de verificación de firmas en línea basado en redes neuronales recurrentes.

El conjunto de experimentos realizado se diseñó para un conjunto de datos formado por firmas realizadas con stylus. Sin embargo, posteriormente estos experimentos se evaluaron sobre un conjunto de firmas realizadas con el dedo. Sumado a la evaluación del rendimiento de los modelos obtenidos sobre firmas realizadas con el dedo, también se realizaron dos experimentos entrenando con este tipo de firmas. El protocolo experimental usado en estos experimentos fue el mismo que el utilizado para los experimentos sobre firmas realizadas con stylus. La única diferencia radica en que para el primer conjunto de experimentos se utilizaron las firmas hechas a stylus de DeepSignDB y en el segundo las firmas realizadas con el dedo.

5.2. Protocolo Experimental

El protocolo experimental se ha diseñado para analizar dos ideas principales:

- Cómo afecta la complejidad de las firmas al rendimiento de un sistema de verificación de firmas en línea.
- Cómo considerar la complejidad de las firmas en el proceso de entrenamiento puede mejorar el rendimiento del sistema de verificación de firmas.

Se han llevado a cabo cinco bloques separados de experimentos. Primero, en el Exp. 1 se analiza el rendimiento de los sistemas DTW y TA-RNN de referencia para cada grupo de complejidad. En el Exp. 2 se entrena desde cero un sistema específico para cada grupo de complejidad utilizando la arquitectura TA-RNN. Luego, en el Exp. 3 se analiza un enfoque similar al Exp. 2 pero modificando la arquitectura del sistema para cada grupo de complejidad. Después, en el Exp. 4 se utiliza un enfoque similar al Exp. 2, pero ajustando cada sistema de grupo de complejidad a partir del sistema de referencia en lugar de entrenar desde cero. Finalmente, en el Exp. 5 seguimos un enfoque diferente, que se basa en el entrenamiento de un solo sistema global, similar al del baseline, pero con un número de sujetos por grupo de complejidad equilibrado.

DeepSignDB ha sido la base de datos utilizada para todos los experimentos realizadas en esta primera fase. Se ha seguido el protocolo experimental de [30], dividiendo los usuarios en dos conjuntos de datos diferentes, uno de desarrollo y otro de evaluación. El conjunto de desarrollo contiene el 70 % de los sujetos, mientras que el otro 30 % es el conjunto de evaluación. Es importante señalar que cada conjunto de datos contiene sujetos diferentes para evitar resultados sesgados.

A partir de los 1084 sujetos originales contenidos en el conjunto de desarrollo, se aplica el detector de complejidad de firmas propuesto en [55] a todas las firmas genuinas. A continuación, se calcula la complejidad media de cada sujeto, teniendo en cuenta la complejidad de todas sus firmas genuinas. Finalmente, se clasifica a cada sujeto en el grupo de complejidad más cercano a su complejidad media. Al final se obtienen 230 sujetos de alta complejidad, 637 sujetos de complejidad media y 217 sujetos de baja complejidad en el conjunto de desarrollo. En este punto algunos sujetos son descartados, aquellos para los que menos de 2/3 de su firma no coincidían con respecto al grupo de complejidad.

El resultado del detector de complejidad en términos de sujetos fue: 373 sujetos de alta complejidad, 756 sujetos de complejidad media y 271 sujetos de baja complejidad. No se consideraron todos los sujetos, ya que sólo se seleccionaron los sujetos con más de un 66 % de confianza en la clasificación. Por ejemplo, un sujeto con una complejidad media de 1,67 se clasifica como de complejidad media al igual que un sujeto con una complejidad media de 2,32. Sin embargo, un sujeto con una complejidad media de 1,50 no se tendrá en cuenta para los experimentos.

Para el entrenamiento del sistema, el conjunto de desarrollo tiene un total de 980 sujetos. Este conjunto se ha dividido a su vez en dos subconjuntos diferentes, uno para el entrenamiento (80 %) y otro para la validación (20 %).

En el caso de firmas realizadas con el dedo a partir de los 102 sujetos originales del conjunto de desarrollo se obtienen 70 usuarios para dicho conjunto: 24 de baja complejidad, 35 de media complejidad y 11 de alta complejidad.

Como el primer objetivo es analizar el rendimiento del sistema para los tres grupos de complejidad por separado, el conjunto de datos de evaluación también se dividió en tres subconjuntos, uno por cada grupo de complejidad. Esto se hizo basándose en el nivel de complejidad de la firma genuina de los pares de firmas comparados, como muestra la Fig. 1. Para proporcionar un rendimiento global del sistema, se juntaron las puntuaciones de los tres conjuntos de evaluación.

Se han considerado dos escenarios de impostores, falsificaciones *skilled* y *random*. En el caso de la falsificación *skilled*, todas las muestras de falsificación disponibles se incluyen en el análisis, mientras que en el caso de la falsificación *random* se incluye una muestra auténtica de cada uno de los sujetos restantes de la misma base de datos. De este modo, los sistemas de verificación se prueban con diferentes tipos de ataques [61]. Cabe destacar que los resultados de evaluación reportados en este trabajo se basan únicamente en comparaciones de firmas uno a uno. Por lo tanto, solo se considera una firma genuina de inscripción por sujeto, lo cual es un caso extremo ya que normalmente se considera más de una firma como inscripción.

Los resultados de la evaluación se presentan como el rendimiento del sistema en términos de curvas DET y de tasa de error igual (EER). Los resultados de la evaluación se obtienen sobre los tres conjuntos de datos de evaluación, uno por clase, así como sobre el conjunto de datos de evaluación completo. De este modo, podemos establecer un análisis comparativo de los resultados obtenidos para cada clase. Se utilizan dos sistemas diferentes como referencia (*baseline*) para comparar los resultados obtenidos en este trabajo. Utilizamos un enfoque tradicional, un sistema basado en DTW, y un sistema de última generación [30].

5.3. Trabajo experimental

5.3.1. Experimentos sobre firmas realizadas con stylus

Exp. 1 - Análisis de la complejidad

En este experimento se lleva a cabo el análisis del rendimiento de evaluación de los dos sistemas de referencia (DTW y TA-RNN) para los tres grupos de complejidad de firma. De esta forma podemos obtener el rendimiento de los sistemas *baseline* sobre los mismos conjuntos de datos de evaluación y realizar un análisis comparativo con los siguientes enfoques propuestos. Las tablas 5.1 y 5.2 muestran el rendimiento del sistema para cada grupo de complejidad para el caso de comparaciones de falsificaciones *skilled* y *random* respectivamente.

En primer lugar, cabe destacar el rendimiento muy superior del sistema TA-RNN en comparación con el enfoque DTW tradicional, en particular para las falsificaciones *skilled*. Tanto en el caso de las falsificaciones *skilled* como en el de las *random*, el mejor rendimiento se obtiene para los sujetos de complejidad media y en ambos sistemas de referencia, con un 9,94 % y un 2,40 % de EER para el sistema DTW, y un 3,32 % de EER y un 1,19 % de EER para el sistema TA-RNN. Esto puede deberse a varios factores. Por un lado, el número de sujetos de complejidad media es el más alto en el conjunto de datos de desarrollo, por lo que es normal que el rendimiento sobre este grupo sea mejor, ya que tiene casi el doble de sujetos que las otras dos clases. Por otro lado, las firmas de complejidad media pueden ser más estables y robustas, logrando una EER más baja. Las firmas de baja complejidad pueden confundirse fácilmente con firmas de sujetos diferentes debido a la baja variabilidad entre clases. Las firmas de alta complejidad del mismo sujeto pueden ser muy diferentes debido a la alta variabilidad intraclase. El segundo mejor rendimiento se consigue con las firmas de alta complejidad, dejando a las firmas de baja complejidad con el peor rendimiento.

En términos generales, el rendimiento global para los sistemas de referencia es de 11,13 % EER para DTW y 4,20 % EER para TA-RNN en falsificaciones *skilled*, y 2,62 % EER para DTW y 1,51 % EER para TA-RNN en falsificaciones *random*. La Fig. 5.1 también muestra la curva DET para la evaluación global del sistema de referencia TA-RNN.

	Baja	Media	Alta	Global
Exp. 1 - Sistema de ref. DTW	13.27	9.94	11.53	11.13
Exp. 1 - Sistema de ref. TA-RNN	5.88	3.32	4.60	4.20
Exp. 2 - 3 Modelos de cero	8.89	3.66	6.31	5.75
Exp. 3 - Arquitectura Adaptada	6.80	4.20	5.63	5.11
Exp. 4 - 3 Modelos Fine Tuning	6.03	3.30	4.60	4.23
Exp. 5 - Modelo Equilibrado	**5.63**	**3.27**	**3.89**	**3.92**

Cuadro 5.1: Comparativa de rendimiento sobre falsificaciones *skilled* en términos de EER.

	Baja	Media	Alta	Global
Exp. 1 - Sistema de ref. DTW	2.62	2.40	2.86	2.62
Exp. 1 - Sistema de ref. TA-RNN	2.08	1.19	1.84	1.51
Exp. 2 - 3 Modelos de cero	3.40	1.24	2.22	2.02
Exp. 3 - Arquitectura Adaptada	1.72	1.45	2.26	1.73
Exp. 4 - 3 Modelos Fine Tuning	2.00	1.17	1.84	1.54
Exp. 5 - Modelo Equilibrado	**1.50**	**1.09**	**1.62**	**1.32**

Cuadro 5.2: Comparativa de rendimiento sobre falsificaciones *random* en términos de EER.

Exp. 2 - Entrenando sistemas adaptados a la complejidad, uno por cada nivel de complejidad

Del Exp. 1 se desprende que el sistema basado en TA-RNN supera claramente al sistema tradicional basado en DTW. Por lo tanto, en los siguientes experimentos el objetivo es explotar la complejidad de la firma para intentar mejorar el ya muy buen rendimiento del sistema TA-RNN. Como primera aproximación para explotar la complejidad de las firmas en un sistema de verificación de firma dinámica, intentamos entrenar tres modelos diferentes desde cero utilizando la arquitectura TA-RNN, uno por grupo de complejidad. Cada modelo se entrenó sólo utilizando comparaciones en las que la firma genuina pertenece a una clase de complejidad específica, pero en las que la firma impostora puede pertenecer a cualquier grupo de complejidad. Es importante mencionar que se utilizaron todos los sujetos disponibles por clase, por lo que a la hora de evaluar los resultados conviene tener esto en cuenta, ya que cada grupo de complejidad tiene diferentes sujetos y, por tanto, diferentes cantidades de datos con los que entrenar. En concreto, se utilizaron 181 sujetos para entrenar el modelo de baja complejidad, 502 sujetos para entrenar el modelo de complejidad media y 184 para entrenar el modelo de alta complejidad. Las tablas 5.1 y 5.2 muestran los resultados obtenidos por cada uno de los modelos entrenados en este experimento siguiendo el mismo protocolo de evaluación que los sistemas de referencia mostrados en el Exp. 1.

Los resultados obtenidos en este experimento nos llevan a una conclusión similar a la anterior en cuanto a los grupos de complejidad, ya que siguen tendencias similares. Sin embargo, los resultados obtenidos con este enfoque compuesto por tres sistemas, uno por grupo de complejidad, son significativamente peores en comparación con los obtenidos por el sistema TA-RNN de referencia. Esto puede deberse a que el sistema de referencia ha sido entrenado con un número mucho mayor de sujetos, y parece que el sistema de referencia puede aprovechar la información proporcionada por la mayor variabilidad de sujetos con los que se ha entrenado y proporcionar mejores resultados. Sólo en el caso de la evaluación de complejidad media, la EER del sistema de referencia y de este enfoque son bastante similares, ya que el sistema de complejidad media se ha entrenado con el mayor número de sujetos.

En términos generales, el rendimiento global para este enfoque es de 5,75 % de EER para falsificaciones *skilled* y de 2,02 % de EER para falsificaciones *random*, lo que también es peor que el rendimiento del sistema de referencia reportado en el Exp. 1. La Fig. 5.1 también muestra las curvas DET con la evaluación global para este enfoque, mostrando el peor rendimiento de todos los experimentos realizados.

Exp. 3 - Adaptando la arquitectura

Este experimento se diseñó para ver si algún cambio en la arquitectura del sistema respecto a la arquitectura original de la TA-RNN podía mejorar su rendimiento frente a un grupo concreto de complejidad. Se entrenaron dos modelos diferentes, uno centrado en firmas de baja

complejidad y el otro centrado en mejorar el rendimiento contra firmas de alta complejidad. Sólo se consideraron estos dos grupos, ya que son las clases en las que se obtuvieron los peores resultados en el Exp. 2.

En el modelo de baja complejidad se eliminó la segunda capa BGRU para crear un modelo más sencillo. Los resultados obtenidos con este sistema fueron realmente malos para todos los tipos de firmas, no sólo para las de baja complejidad. Por otro lado, en el modelo enfocado a la alta complejidad se añadió una tercera capa oculta BGRU antes de la capa feed forward. Este nuevo modelo entrenado sólo con sujetos de alta complejidad resultó dar resultados razonables en general.

La segunda arquitectura propuesta, sólo entrenada con sujetos de alta complejidad, obtuvo mejores resultados. Las tablas 5.1 y 5.2 muestran el rendimiento obtenido por este sistema. Como podemos ver los resultados obtenidos por el modelo de alta complejidad mejoran los resultados obtenidos por los modelos del Exp. 2, tanto en alta como en baja complejidad. Esto también se puede ver en términos generales en las curvas DET mostradas en la Fig. 5.1.

A pesar de no ser capaz de mejorar los resultados del sistema de referencia, esta nueva arquitectura es prometedora ya que consigue resultados más cercanos a los del sistema de referencia a pesar de estar entrenado con un número mucho menor de sujetos.

Exp. 4 - Fine tuning

Además de los cambios en la arquitectura, en este experimento también entrenamos tres sistemas similares al Exp. 2, pero en lugar de entrenar desde cero, aplicamos el ajuste fino del sistema de referencia del Exp. 1 a cada sistema del grupo de complejidad.

Este experimento se diseñó para averiguar si el rendimiento del sistema podía mejorarse para un grupo de complejidad concreto llevando a cabo un ajuste fino del modelo de referencia utilizando únicamente los sujetos de ese grupo de complejidad. Para cada complejidad, se realizaron tres tipos de fine tuning: entrenar sólo la última capa (totalmente conectada), entrenar las dos últimas capas (segunda BGRU y capas totalmente conectadas) y entrenar todas las capas (las dos BGRU y las capas totalmente conectadas). Por lo tanto, se entrenaron tres modelos diferentes por grupo de complejidad.

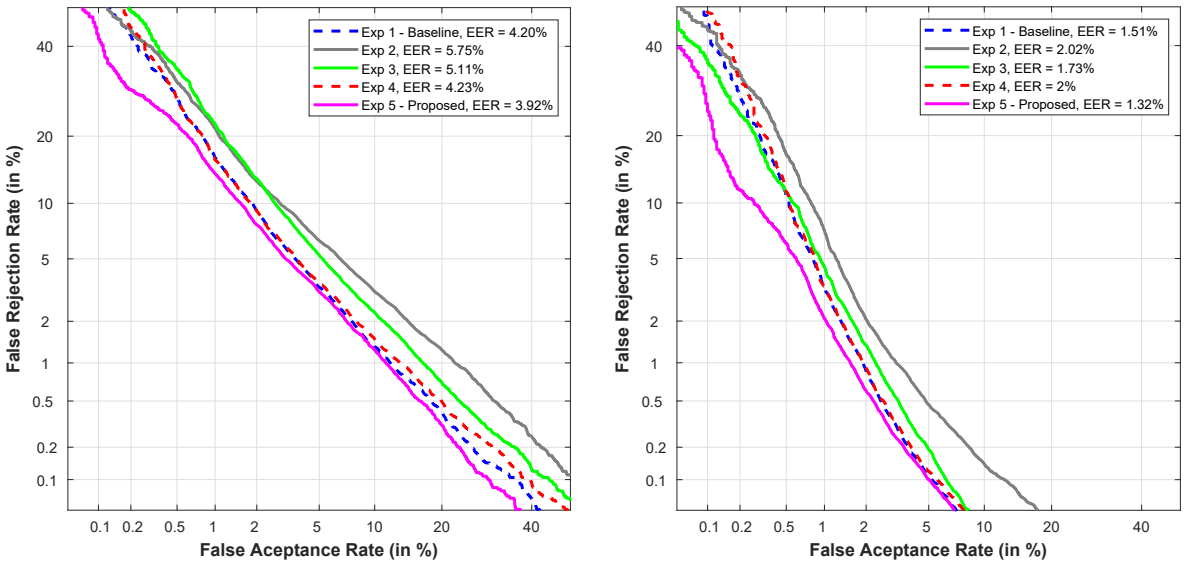
Se obtuvieron resultados similares con las distintas estrategias de ajuste. Los mejores resultados se obtuvieron entrenando sólo la capa totalmente conectada. Las tablas 5.1 y 5.2 muestran los resultados obtenidos por los tres modelos que aplican esta estrategia.

Podemos ver que apenas hay diferencia con el rendimiento del sistema de referencia. En el caso de las firmas de baja complejidad, hay una pequeña caída en el rendimiento con respecto a las falsificaciones *skilled* y una sutil mejora con respecto a las falsificaciones *random*. Si se realiza una puesta a punto con más sujetos de baja y alta complejidad se podrían obtener mejores resultados. Además, observando las curvas de DET en la Fig. 5.1 podemos ver cómo el rendimiento conseguido con este enfoque es muy similar al conseguido por el sistema de referencia TA-RNN en términos generales.

Exp. 5 - Entrenando un sistema global con clases equilibradas

Después de haber analizado muchos enfoques diferentes entrenando sistemas específicos por grupo de complejidad, decidimos entrenar un solo sistema pero utilizando un número equilibrado de sujetos (y muestras) por grupo de complejidad.

Se entrenó un modelo utilizando la arquitectura TA-RNN. Para entrenar el sistema se utilizaron 181 sujetos por clase, en total 543. De esta forma se utilizaron todos los sujetos posibles



(a) Falsificaciones Skilled

(b) Falsificaciones Random

Figura 5.1: Resultado del rendimiento de los experimentos sobre el conjunto de evaluación de DeepSignDB.

manteniendo un número equilibrado de sujetos entre los diferentes grupos de complejidad. Durante el proceso de entrenamiento se siguieron diferentes estrategias, es decir, entrenamiento desde cero y ajuste fino de una, dos y las tres capas de la arquitectura TA-RNN. Los mejores resultados se obtuvieron al entrenar el modelo desde cero.

Las tablas 5.1 y 5.2 muestran el rendimiento de este enfoque. A pesar de haber sido entrenado con un número reducido de sujetos en comparación con el sistema de referencia (543 frente a 1084 sujetos), observamos que el sistema consigue un mejor rendimiento para todos los tipos de complejidad y también en la evaluación general.

Analizando el caso de las falsificaciones *skilled*, la mejora del rendimiento para la complejidad baja y media es pequeña. Sin embargo, para la complejidad alta hay una mejora relativa del 15,44 % EER en comparación con el sistema de referencia TA-RNN. En cuanto al rendimiento global, la mejora relativa conseguida con el sistema propuesto es de un 6,67 % de EER en falsificaciones *skilled*.

Analizando el caso de las falsificaciones *random*, la mejora conseguida es más significativa en comparación con el caso de las falsificaciones *skilled*. En este caso, se consigue una mejora relativa muy significativa del 27,89 % de EER para el caso de baja complejidad con un 1,50 % de EER. Las mejoras relativas para la complejidad media y alta son del 8,40 % y del 11,96 % de EER respectivamente. En cuanto al rendimiento global, la mejora relativa conseguida con el sistema propuesto es del 12,58 % de EER para falsificaciones *random*.

Por último, en la Fig. 5.1 se muestran las curvas de DET para el sistema propuesto, logrando el mejor rendimiento del sistema en todos los puntos de funcionamiento de la tasa de falsos rechazos (FRR) y la tasa de falsa aceptación (FAR), en comparación con todos los experimentos realizados y los sistemas de referencia. Esto pone de manifiesto la importancia de contar con clases equilibradas en la fase de entrenamiento. Analizando los resultados obtenidos podemos concluir que es más importante tener un número equilibrado de sujetos por clase que un alto número de sujetos desequilibrados.

Análisis del rendimiento sobre firmas realizadas con el dedo

Una vez finalizados los experimentos sobre el conjunto de datos de firmas realizadas con stylus, se procedió a evaluar los modelos resultantes de dichos experimentos sobre el conjunto de datos conformado por las firmas realizadas con el dedo. Para este análisis se ha usado el conjunto de evaluación de las firmas disponibles en las bases de datos e-BioSign DS1 y e-BioSign DS2.

En las tablas 5.3 y 5.4 se muestra el rendimiento en término de EER de los modelos obtenidos en los experimentos anteriores sobre el conjunto de evaluación de firmas realizadas con el dedo. El formato de estas tablas es el mismo que el de las anteriores, mostrando el rendimiento sobre falsificaciones *skilled* en la primera tabla y el rendimiento sobre falsificaciones *random* en la segunda.

	Baja	Media	Alta	Global
Exp. 1 - Sistema de ref. DTW	18.04	15.02	16.74	16.32
Exp. 1 - Sistema de ref. TA-RNN	15.02	12.98	14.02	13.84
Exp. 2 - 3 Modelos de cero	20.92	14.05	20.64	17.44
Exp. 3 - Arquitectura Adaptada	21.01	18.51	19.70	19.55
Exp. 4 - 3 Modelos Fine Tuning	15.10	13.33	14.21	14.07
Exp. 5 - Modelo Equilibrado	**15.54**	**11.55**	**18.37**	**14.00**

Cuadro 5.3: Comparativa de rendimiento sobre falsificaciones *skilled* en términos de EER para firmas realizadas con el dedo.

	Baja	Media	Alta	Global
Exp. 1 - Sistema de ref. DTW	3.76	3.78	4.16	3.83
Exp. 1 - Sistema de ref. TA-RNN	1.56	2.08	1.14	1.75
Exp. 2 - 3 Modelos de cero	1.96	1.70	3.98	2.15
Exp. 3 - Arquitectura Adaptada	3.91	4.2	3.98	4.07
Exp. 4 - 3 Modelos Fine Tuning	1.58	2.06	1.14	1.75
Exp. 5 - Modelo Equilibrado	**2.08**	**2.50**	**2.17**	**2.30**

Cuadro 5.4: Comparativa de rendimiento sobre falsificaciones *random* en términos de EER para firmas realizadas con el dedo.

En un primer vistazo lo que podemos observar es la bajada de rendimiento que obtienen todos los sistemas para las firmas realizadas con el dedo. Este efecto era esperable ya que ninguno de estos sistemas fue entrenado con este tipo de firmas. La verificación sobre este tipo de firmas suele ser más compleja debido a la falta de precisión con la que contamos al usar el dedo como instrumento de firma, aumentando la variabilidad intra-clase y disminuyendo la estabilidad.

Atendiendo al rendimiento para cada grupo de complejidad, la conclusión que podemos extraer de estos resultados es bastante similar a la obtenida en los experimentos anteriores.

En ambos sistemas baseline obtenemos el mejor rendimiento para las firmas de clase media, seguido por las firmas de clase alta y en último lugar las de clase baja. La hipótesis expuesta anteriormente también puede explicar estos resultados. Las firmas de clase media al ser el grupo con mayor número de usuarios y un tipo de firma con mayor estabilidad son las que mejores resultados obtienen. Las firmas de baja complejidad debido a su simpleza son fácilmente imitables, lastrando el rendimiento del sistema sobre falsificaciones *skilled*, y poco diferenciables de firmas simples de otros usuarios, aumentando el error frente a falsificaciones *random*. Las firmas de alta complejidad si que son más distinguibles frente a falsificaciones y firmas de otros usuarios, pero la alta variabilidad intra-clase derivada de su complejidad hacen que el sistema baje el rendimiento.

La diferencia de rendimiento entre el Exp. 1 y el Exp. 2 sobre las firmas realizadas con el dedo es similar al obtenido sobre las firmas realizadas con stylus. Por lo tanto, podemos llegar a la misma conclusión: los modelos del Exp. 2 obtienen peor resultado que el sistema TA-RNN debido al uso de un menor número de usuarios en su entrenamiento.

Curiosamente, en este caso el modelo con la arquitectura adaptada para la complejidad alta obtiene peor resultado que los modelos del Exp 2. Esto puede ser debido a la menor calidad de las firmas realizadas con el dedo, que hace que no se capture todo el detalle de las firmas de alta complejidad, empeorando el rendimiento del sistema.

En cuanto al resultado obtenido sobre los modelos del Exp. 4 no hay ninguna conclusión adicional, estos modelos, al igual que sobre firmas realizadas con stylus, obtienen un rendimiento muy similar al del sistema TA-RNN.

Finalmente, la diferencia la encontramos en el rendimiento del modelo del Exp. 5. En este caso, no solo no se mejoran los resultados del sistema TA-RNN sino que estos empeoran. Además, podemos observar que la mayor pérdida de rendimiento ocurre sobre las firmas de alta complejidad, para las que se obtiene el peor rendimiento de todos los grupos de complejidad. De hecho, estos malos resultados son los que lastran el rendimiento global, ya que para los grupos de media y baja complejidad el rendimiento es similar al del sistema TA-RNN; con una mejoría para las firmas de media complejidad en falsificaciones *skilled* y un leve empeoramiento para falsificaciones *random*.

5.3.2. Experimentos sobre firmas realizadas con el dedo

Una vez analizado los efectos de la complejidad sobre las firmas realizadas con el dedo, se hicieron dos experimentos con el fin de aprovechar estos efectos para mejorar al eficiencia de los sistemas de verificación sobre este tipo de firmas.

Como se ha mencionado antes, el protocolo experimental que se ha seguido a la hora de llevara cabo estos experimentos ha sido el mismo que el utilizado para los experimentos sobre firmas realizadas con stylus, simplemente variando los datos utilizados.

Se han llevado a cabo dos experimentos principales, el primero de ellos consiste en entrenar un modelo equilibrado de cero mientras que el segundo consiste en hacer fine-tuning del modelo obtenido en el Exp. 5.

Exp. 6 - Entrenando modelos de cero

En primer lugar, con el objetivo de hacer una comparación parecida a la de los experimentos anteriores, se entrenó un modelo usando todo el conjunto de entrenamiento de firmas hechas con el dedo y haciendo uso de la arquitectura TA-RNN. De esta manera, se consigue un sistema que no tiene en cuenta la información de complejidad. Para el entrenamiento de dicho modelo se han usado los 70 usuarios del conjunto de entrenamiento.

Una vez entrenado el modelo genérico que no utiliza la información de complejidad, se entrenó un modelo con un número de usuarios equilibrado en base a su complejidad. En total se usaron 33 usuarios, 11 por cada grupo de complejidad. Por lo tanto, prácticamente la mitad de usuarios pero teniendo en cuenta la complejidad de los mismos. De esta forma, podemos ver la importancia de la información que nos aporta la complejidad de las firmas mediante la comparación del rendimiento de ambos sistemas.

En la tabla 5.5 se muestra la comparación de rendimiento de los dos sistemas en términos de EER para cada grupo de complejidad y para dos tipos de falsificaciones (*skilled* y *random*).

		Baja	Media	Alta	Global
Exp. 6 - Modelo Genérico	Skilled	15.10	9.40	14.96	12.23
	Random	**3.60**	**2.77**	**2.84**	**3.07**
Exp. 6 - Modelo Equilibrado	Skilled	**13.37**	**10.83**	**13.07**	**12.05**
	Random	5.09	4.63	4.85	4.82

Cuadro 5.5: Comparativa de rendimiento sobre falsificaciones *skilled* y *random* en términos de EER.

Como podemos ver, los resultados del modelo equilibrado superan a los del genérico en rendimiento sobre las falsificaciones *skilled*, a pesar de ser entrenado con un número muy inferior. Esto nos denota la importancia que la información de complejidad adquiere en la detección de este tipo de falsificaciones. En el único caso en el que el modelo genérico obtiene un mejor resultado es en el de las firmas de media complejidad. Seguramente este resultado sea debido a la mayor cantidad de usuarios de clase media usados para entrenar el modelo genérico (35 usuarios), que triplican los usados para entrenar el modelo equilibrado (11 usuarios).

Por otro lado, el modelo genérico obtiene mejores resultados frente a falsificaciones *random*. En este caso prima el número de usuarios de entrenamiento frente a la información de complejidad. Esta hipótesis tiene sentido debido a que este tipo de falsificaciones son más fáciles de detectar, lo que hace que no se necesite tanta información para su detección.

Exp. 7 - Aplicando Fine Tuning

Tras el análisis realizado en el experimento anterior sobre los efectos de la complejidad en sistemas de verificación de firma dinámica realizada con el dedo, se procedió a intentar mejorar estos sistemas a partir del modelo equilibrado obtenido en los experimentos realizados sobre firma hecha con stylus. A pesar de no ser el modelo que mejores resultados arroja para firmas realizadas con dedo, se eligió debido a sus buenos resultados sobre firmas realizadas con stylus. Al no haber entrenado ninguno de los sistemas del primer conjunto de experimentos con firmas realizadas con dedo, la comparativa no es directa y los resultados sobre stylus nos permiten diferenciar mejor que tipo de sistema es más eficaz.

Se realizó un fine-tuning del modelo obtenido en el Exp. 5 con el conjunto de desarrollo de dedo. Para realizar el fine-tuning no se congeló ninguna de las capas, se reentrenó el modelo completo. Se siguió esta metodología debido a que tanto la capa de clasificación como la de extracción de características tenían que adaptarse al nuevo tipo de firmas.

Inicialmente, el reentrenamiento del modelo se realizó con un número de usuarios equilibrado en base a su complejidad, manteniendo de esta forma la información de complejidad. Después, se llevo a cabo el fine tuning del modelo con todas las firmas disponibles en el conjunto de entrenamiento de dedo, obviando de esta forma la información de complejidad, pero aumentando considerablemente el número de usuarios.

En las tablas 5.6 y 5.7 podemos observar la comparativa de resultados en términos de EER entre los dos modelos entrenados en este experimento y el modelo obtenido en el Exp. 5.

Claramente, vemos que el entrenamiento con firmas realizadas con el dedo mejora notablemente el rendimiento del sistema. La diferencia existente entre los dos tipos de firmas es notoria, teniendo las firmas a dedo una estabilidad menor y una variabilidad intra-clase mayor, mientras que la variabilidad inter-clase se reduce. De esta forma, el sistema se adapta a este nuevo tipo de firmas tras tener acceso a la información que aportan las mismas consiguiendo así mejorar su rendimiento.

	Baja	Media	Alta	Global
Exp. 1 - Sistema de ref. TA-RNN	15.02	12.98	14.02	13.84
Exp. 5 - Modelo Equilibrado	15.54	11.55	18.37	14.00
Exp. 7 - Fine Tuning Equilibrado	10.59	7.32	11.74	9.14
Exp. 7 - Fine Tuning Genérico	**10.85**	**6.19**	**8.14**	**8.09**

Cuadro 5.6: Comparativa de rendimiento sobre falsificaciones *skilled* en términos de EER.

	Baja	Media	Alta	Global
Exp. 1 - Sistema de ref. TA-RNN	1.56	2.08	1.14	1.75
Exp. 5 - Modelo Equilibrado	2.08	2.50	2.17	2.30
Exp. 7 - Fine Tuning Equilibrado	2.02	2.06	3.14	2.22
Exp. 7 - Fine Tuning Genérico	**1.70**	**1.48**	**1.10**	**1.50**

Cuadro 5.7: Comparativa de rendimiento sobre falsificaciones *random* en términos de EER.

Cabe destacar que ambos sistemas consiguen una gran mejora de rendimiento a pesar del pequeño número de usuarios disponibles para el entrenamiento. El mejor resultado se obtiene para el modelo genérico, obteniendo un 8.09 % de EER sobre falsificaciones *skilled* y un 1.50 % de EER sobre *random*. Esto es debido a la carencia de información sobre firmas a dedo que tiene el sistema, haciendo que cuanto mayor sea el número de usuarios utilizados para el entrenamiento mejor rendimiento se obtenga. En este caso, la información de complejidad ya ha sido previamente aprendida por el sistema en el Exp. 5, por lo que el entrenamiento equilibrado no añade información adicional.

Podemos concluir que es preferible aportar un mayor número de muestras de entrenamiento de forma que el sistema pueda adaptarse a las firmas realizadas con dedo, en vez de disponer de menos muestras con el mismo número de usuarios por grupo de complejidad. Sin embargo, como demuestra el Exp. 6, si nuestro sistema ya está adaptado a este tipo de firmas, tener en cuenta la complejidad nos aporta una mejora de rendimiento.

6

Resultados del desarrollo de sistemas de firma estática y fusión de sistemas

6.1. Desarrollo de sistemas de firma estática

En este segundo conjunto de experimentos se busca mejorar sistemas de verificación de firma estática con la finalidad de combinarlos con sistemas de verificación de firma dinámica. Por lo tanto, debido a que el desarrollo de estos sistemas está enfocado a su posterior fusión con sistemas dinámicos, sólo se usarán firmas estáticas sintéticas para el entrenamiento y evaluación de los mismos.

Todos los desarrollos llevados a cabo han partido de la capa de extracción de características Signet. Esta elección ha sido debida a los buenos resultados que el sistema propuesto en [41] arrojaba. A pesar de que estos resultados eran obtenidos por el sistema completo, no sólo por la red Signet, esta red juega un papel fundamental en el rendimiento del sistema realizando la extracción de características. A partir de estas características se intentarán buscar métodos alternativos de clasificación de las mismas.

6.1.1. Protocolo Experimental

El protocolo experimental seguido en esta segunda fase de experimentos ha sido muy similar al utilizado en la primera parte de este trabajo. Esto es debido a que, aunque esta parte se centrará en la mejora de sistemas de verificación de firma estática, el objetivo final es fusionar ambos sistemas y es necesario disponer de un marco común de entrenamiento y evaluación.

Se han llevado a cabo tres tipos de experimentos. Primero se realizó una evaluación del rendimiento que se podía obtener usando únicamente las características extraídas por Signet. Después, se desarrolló una capa de clasificación basada en redes neuronales en configuración siamesa con el objetivo de mejorar el rendimiento del sistema. Finalmente, se llevó a cabo el refinamiento de las características obtenidas por Signet mediante triplet loss.

La base de datos utilizada ha sido DeepSignDB. A partir de los datos dinámicos de las firmas se ha realizado un proceso de generación de una base de datos estática sintética. Por lo tanto, para el entrenamiento y evaluación de los sistemas se ha utilizado las mismas firmas que en la primera parte de este trabajo, pero en versión estática sintética.

La división de los datos ha seguido el mismo patrón que en el anterior bloque de experimentos. Se ha dividido la base de datos en un conjunto de desarrollo (70 %) y un conjunto de evaluación (30 %). A su vez el conjunto de desarrollo se ha divido en entrenamiento (80 %) y validación (20 %). En esta ocasión no se ha tenido en cuenta la información de complejidad por lo que no se ha realizado la división por grupos de complejidad. Tampoco se han usado firmas realizadas con el dedo en estos experimentos.

Al igual que en la primera parte de este trabajo, se han considerado dos escenarios de impostores, falsificaciones *skilled* y falsificaciones *random*. Todas las falsificaciones *skilled* disponibles por usuario se han incluido en el análisis mientras que como falsificaciones *random* se ha tenido en un cuenta una muestra genuina de cada uno de los sujetos restantes de la misma base de datos.

Los resultados de la evaluación se presentan como el rendimiento del sistema en términos de EER. Estos resultados se obtienen sobre el conjunto completo de evaluación de DeepSign. Como sistema de referencia se utiliza la red de extracción de características Signet propuesta en [41], la cuál fue entrenada con firmas estáticas de la base de datos GPDS. La distancia euclídea entre estas características será, por lo tanto, el sistema con el que se comparen los resultados obtenidos en este trabajo.

6.1.2. Trabajo Experimental

Evaluación del sistema de referencia

Como hemos comentado antes, inicialmente se realizó una evaluación del rendimiento que podía obtener la red Signet. Para ello, se utilizó la distancia euclídea entre las características extraídas por dicha red como método de comparación. Sumado a la evaluación del modelo propuesto en [41], se evaluó un nuevo modelo entrenado con la base de datos DeepSign. Dicho sistema cuenta con la misma arquitectura, lo único que varían son los datos de entrenamiento.

Las tablas 6.1 y 6.2 describen el rendimiento de las diferentes aproximaciones evaluadas en términos de EER para falsificaciones *skilled* y *random* sobre la base de datos DeepSignDB. La tabla 6.1 muestra los resultados para comparaciones con una muestra genuina (1vs1) y la tabla 6.2 muestra los resultados obtenidos al realizar la comparación de una muestra de test con cuatro firmas genuinas del mismo usuario (4vs1).

	Skilled	Random
Sistema de ref. Signet	27.58	10.58
Sistema de ref. entrenado con DeepSign	21.50	3.97
Arquitectura Siamesa	20.9	8.30
Capa de Clasificación Neuronal	**20.02**	**4.54**
Triplet Loss	26.54	9.50
Triplet Loss Fine Tuning	21.21	7.64

Cuadro 6.1: Evaluación de resultados 1vs1: Comparativa de rendimiento sobre falsificaciones *skilled* y *random* en términos de EER.

El sistema Signet no consigue buenos resultados sobre falsificaciones *skilled*, donde tiene un 27.58 % de EER para comparaciones 1vs1. Al evaluar su rendimiento sobre falsificaciones *random* vemos que los resultados obtenidos son bastante mejores (10.58 %), pero todavía distantes del estado del arte. La evaluación para comparaciones 4vs1 nos da cierta mejora derivada de un mayor número de comparaciones. Sin embargo, seguimos obteniendo resultados insuficientes para un sistema completo, pero aceptables si tenemos en cuenta que el único proceso realizado es la

	Skilled	Random
Sistema de ref. Signet	23.40	7.41
Sistema de ref. entrenado con DeepSign	18.54	2.93
Arquitectura Siamesa	17.62	5.99
Capa de Clasificación Neuronal	**16.92**	**3.28**
Triplet Loss	24.11	7.37
Triplet Loss Fine Tuning	18.33	5.62

Cuadro 6.2: Evaluación de resultados 4vs1: Comparativa de rendimiento sobre falsificaciones *skilled* y *random* en términos de EER.

extracción de características. Por ello, podemos deducir que estas características son robustas y una buena base para una capa de comparación posterior.

Por otra parte, el modelo entrenado con el conjunto de datos DeepSign consigue una mejora significativa respecto al modelo anterior. Esta mejora probablemente sea debida a que DeepSign es una base de datos con un número elevado de usuarios cuyas firmas han sido obtenidas en escenarios diversos. Esto permite que los modelos entrenados con ella sean robustos frente a escenarios de interoperabilidad y consigan una mejor generalización. Este modelo obtiene un 21.50 % y 3.97 % para comparaciones *skilled* y *random*, respectivamente, en el escenario 1vs1; y 18.54 % y 2.93 % para el escenario 4vs1. En este caso, la mejora obtenida nos da muy buenos resultados de rendimiento sobre falsificaciones *random*, aunque para falsificaciones *skilled* los resultados todavía no son óptimos.

Entrenamiento con arquitectura siamesa

Una vez evaluado el rendimiento de las características extraídas por la red Signet, se procedió a la búsqueda de nuevas aproximaciones para la comparación de estas características. La primera aproximación evaluada fue el entrenamiento del sistema en configuración siamesa añadiendo una capa de comparación conformada por una red neuronal (*Arquitectura Siamesa* en la tabla 6.1). Se decidió evaluar esta aproximación debido al buen resultado que obtiene en sistemas de verificación de firma dinámica.

Como capa de clasificación se evaluaron distintas configuraciones. Se hicieron pruebas variando el número de capas, el número de neuronas por capa y el dropout existente entre una capa y la siguiente. Finalmente, los mejores resultados fueron obtenidos para un modelo de tres capas con 75, 50 y 1 neuronas, respectivamente, y un dropout de 0.3.

Esta configuración resultó en una pequeña mejora de rendimiento sobre las falsificaciones *skilled*, obteniendo un 20.9 % de EER para comparaciones 1vs1 y un 17.62 % para comparaciones 4vs1. Sin embargo, sobre falsificaciones *random* encontramos una pérdida de rendimiento situando el EER en 8.30 % para comparaciones 1vs1 y en 5.99 % para comparaciones 4vs1.

Tras los resultados obtenidos por el sistema siamés, se decidió comprobar si la falta de eficacia de dicho sistema era debido a la capa de clasificación o al entrenamiento siamés, el cuál empeoraba las características extraídas. Para ello, se entrenó un nuevo modelo únicamente consistente en la capa de clasificación del sistema siamés (*Capa de Clasificación Neuronal*.[en] la tabla 6.1), la cuál tiene como entrada las características extraídas por el modelo de referencia entrenado con DeepSign.

Este sistema consigue sobrepasar en rendimiento al anterior, indicándonos por tanto que el entrenamiento siamés no era adecuado para la red Signet. En comparación al sistema de referencia conseguimos de nuevo una mejora en el rendimiento sobre falsificaciones *skilled*, en concreto 20.02 % para el caso 1vs1 y 16.92 % para el caso 4vs1, siendo este resultado el más llamativo. Pero

no sólo mejora el rendimiento sobre falsificaciones *skilled*. En esta ocasión, los resultados obtenidos sobre falsificaciones *random* son muy próximos a los del sistema de referencia, obteniendo un 4.54 % y 3.28 % para los casos 1vs1 y 4vs1.

Podemos concluir, por lo tanto, que el entrenamiento de la red Signet en configuración siamesa no consigue buenos resultados y que el uso de una capa de comparación conformada por una red neuronal, compuesta por capas fully-connected, supone una mejora de rendimiento sobre falsificaciones *skilled*.

Triplet Loss

La segunda aproximación que se planteó fue realizar el entrenamiento de la red Signet usando la función de pérdidas triplet loss (*Triplet Loss* en la tabla 6.1). El objetivo de dicho experimento fue ver si triplet loss era un método de entrenamiento adecuado. Una vez entrenado este modelo se extrajeron las características obtenidas por el mismo y se evaluaron usando como método de comparación la distancia euclídea, como se hizo con el sistema de referencia.

Los resultados obtenidos por este método no fueron satisfactorios, obteniendo rendimientos de entorno al 30 % de EER para ambos tipos de falsificaciones. Estos resultados no se muestran en las tablas 6.1 y 6.2.

Una vez descartado el entrenamiento mediante triplet loss, se decidió comprobar si se podía obtener algún beneficio de dicha función de pérdidas. Para ello se procedió a entrenar una red de extracción de características adicional mediante triplet loss. Esta red tiene como entrada las características extraídas por la red Signet, entrenada con DeepSign, y como salida un conjunto de características menor, las cuáles contienen la información más relevante. Se evaluaron distintas configuraciones para esta capa variando, al igual que con la capa de comparación del apartado anterior, el número de capas neuronales, el número de neuronas y el dropout entre capas. El modelo que mejor resultados arrojó fue el compuesto por dos capas neuronales de 40 neuronas usando un dropout de 0.2.

Este nuevo conjunto de características fue evaluado al igual que el resto de características, usando la distancia euclídea como método de comparación. Los resultados obtenidos denotaron que las características obtenidas extraían peor información que las características extraídas por Signet, no consiguiendo superarlas en rendimiento en ningún escenario.

Finalmente, el último experimento que se llevó a cabo consistió en realizar un fine-tuning a la capa de extracción de características de la arquitectura siamesa (*Triplet Loss Fine Tuning* en la tabla 6.1). Para ello, se extrajeron los pesos de dicha capa y se reentrenaron usando triplet loss. Una vez reentrenada, los nuevos pesos se introdujeron en la red siamesa y se evaluó su rendimiento.

El resultado de esta aproximación fue una mejora de rendimiento, respecto al sistema de arquitectura siamesa, sobre falsificaciones *random*, obteniendo un 7.64 % de EER para el caso 1vs1 y un 5.62 % para el caso 4vs1. Sin embargo, esta mejora vino acompañada de una leve pérdida de rendimiento sobre las falsificaciones *skilled*, con un EER de 21.21 % para las comparaciones 1vs1 y un EER de 18.33 % para las comparaciones 4vs1. Vemos por lo tanto, que triplet loss consigue mejorar el rendimiento sobre falsificaciones *random* a costa de disminuirlo ligeramente sobre falsificaciones *skilled*.

Si bien no podemos concluir en que triplet loss sea un método de entrenamiento poco eficaz para sistemas de verificación de firma estática, si podemos inferir que no es adecuado para los modelos propuestos en este trabajo.

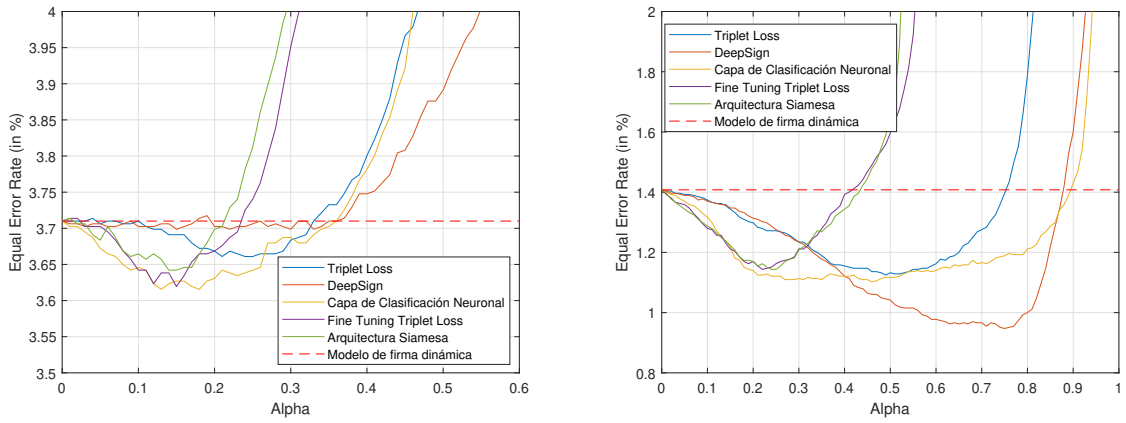
6.2. Fusión de sistemas

Como última fase de este trabajo se procederá a la fusión de sistemas estáticos y dinámicos. El sistema dinámico escogido para la fusión ha sido el modelo equilibrado obtenido en el Exp. 5 del primer conjunto de experimentos expuestos en este trabajo, debido a que es el que mejor resultados obtiene. Se ha procedido a fusionar este sistema con los cinco sistemas descritos en la sección 6.1 del capítulo 6: la red Signet entrenada con DeepSign, el sistema de arquitectura siamesa, el sistema que compara las características de Signet con una red neuronal, el sistema de extracción de características triplet loss y el sistema de arquitectura siamesa reentrenado con triplet loss.

La fusión ha sido realizada a nivel de puntuación, por lo tanto, ha consistido en la combinación de las puntuaciones del sistema dinámico y uno de los sistemas estáticos. Dicha combinación se ha realizado mediante una media ponderada en la cuál un factor α determina el porcentaje atribuido a cada sistema, teniendo los resultados del sistema dinámico con $\alpha = 0$ y los resultados del sistema estático con $\alpha = 1$. Se ha ido variando dicho porcentaje con el fin de evaluar cual es el punto óptimo de fusión. En la ecuación 5.1 vemos la fórmula usada para la fusión.

$$P_{comb} = \frac{P_{estatica} * \alpha + P_{dinamica} * (1 - \alpha)}{2} \tag{6.1}$$

Los sistemas de arquitectura siamesa y el sistema con una capa de comparación formada por una red neuronal devuelven como resultado de la verificación un número entre 0 (genuino) y 1 (impostor), de forma que se puede hacer una combinación directa. Para los sistemas que hacen uso de la distancia euclídea como método de comparación se ha procedido a la normalización de los resultados mediante el método min-max, de forma que podemos realizar la fusión.



(a) Falsificaciones Skilled

(b) Falsificaciones Random

Figura 6.1: Resultados del rendimiento de la fusión de sistemas sobre el conjunto de datos de evaluación DeepSignDB para comparaciones 1vs1. El parámetro α indica el grado de fusión de los sistemas, teniendo para $\alpha = 0$ los resultados del sistema dinámico y para $\alpha = 1$ los resultados del sistema estático.

En la figura 6.1 podemos observar los resultados de la combinación de sistemas en forma de curvas en las que cada punto se corresponde con el rendimiento del sistema combinado en términos de EER para un valor de α determinado. Estos resultados se han obtenido para comparaciones 1vs1. Podemos ver que todas las combinaciones evaluadas proporcionan una mejora de rendimiento los dos tipos de impostores. La tabla 6.3 muestra los resultados obtenidos por cada sistema para un punto global óptimo de fusión ($\alpha = 0{,}2$). Se ha escogido dicho punto al ser el punto intermedio entre los puntos óptimos de fusión para falsificaciones *skilled* y *random*.

	Skilled	Random
Modelo de firma dinámica	3.72	1.39
DeepSign	3.72	1.32
Arquitectura Siamesa	3.68	1.17
Capa de Clasificación Neuronal	**3.63**	**1.14**
Triplet Loss	3.67	1.30
Triplet Loss Fine Tuning	3.66	1.16

Cuadro 6.3: Resultados del rendimiento de la fusión de sistemas para un punto óptimo de fusión ($\alpha = 0{,}2$).

En el escenario de impostores *skilled*, el punto óptimo de fusión se obtiene para un valor de α de 0.15, a excepción del sistema Triplet Loss cuyo punto óptimo se obtiene entorno al 0.25. Todas las combinaciones de sistemas consiguen una leve mejora de rendimiento respecto al sistema dinámico. Considerando los puntos óptimos de fusión, podemos deducir que los sistemas estáticos aportan poca información añadida, por lo que asignando más peso al sistema dinámico se obtienen mejores resultados.

En el escenario de impostores *random*, el punto óptimo de fusión se obtiene para un valor de α de 0.23, con la excepción del sistema Triplet Loss, cuyo punto óptimo es 0.5, y el sistema DeepSign, cuyo punto óptimo es 0.75. En esta ocasión la mejora de rendimiento lograda es sustancialmente superior, consiguiendo alcanzar un EER de 0.88 %. En este escenario, vemos que los sistemas estáticos continúan aportando menos información que el sistema dinámico.

Sin embargo, hay dos excepciones, el sistema Triplet Loss y el sistema DeepSign. Ambos sistemas utilizan a distancia euclídea como etapa de comparación, ya que ambos sistemas extraen características. Estos sistemas si aportan información relevante y por ello se obtienen mejor resultados al otorgarles más peso en la combinación. En concreto, el sistema DeepSign es el que mejor rendimiento obtiene y el que tiene un punto óptimo más cercano a 1.

En conclusión, estos resultados nos indican que los sistemas estáticos y dinámicos capturan información diferente y que la fusión de sistemas nos permite aprovechar todos los datos disponibles para lograr un mejor rendimiento. Si bien, no se ha podido aprovechar el potencial de la fusión usando sistemas estáticos más eficaces, estos resultados nos permiten demostrar los beneficios de la fusión de sistemas. Asimismo, los resultados obtenidos nos confirman que el mejor sistema estático desarrollado ha sido el sistema con la capa de clasificación neuronal y que las características extraídas por DeepSign aportan más información que las extraídas por Triplet Loss.

7

Conclusiones y trabajo futuro

Este trabajo ha sido realizado en tres fases diferentes, siendo las dos primeras independientes entre sí, pero enfocadas a un mismo objetivo llevado a cabo en la último fase.

Inicialmente, se ha propuesto un análisis en profundidad sobre cómo afecta la complejidad de las firmas al rendimiento de un sistema de verificación de firma dinámica. En particular, se ha realizado un análisis del rendimiento del sistema para tres grupos de sujetos en cuanto a la complejidad de su firma: sujetos con una complejidad de firma baja, media y alta. En general, basándonos en dos sistemas de referencia (DTW y TA-RNN), los sujetos con una complejidad de firma media consiguen el mejor rendimiento del sistema tanto para las comparaciones cualificadas como para las aleatorias, seguidos por los usuarios de alta complejidad y, por último, los de baja complejidad. La razón de esto podría ser que las firmas de baja complejidad pueden tener una baja variabilidad entre clases, mientras que las firmas de alta complejidad pueden tener una alta variabilidad intra-clase. Es probable que las firmas de complejidad media tengan una mayor variabilidad inter-clase y una menor variabilidad intra-clase en comparación con las de baja y alta complejidad.

Esta información podría tenerse en cuenta para advertir a los usuarios con firmas de baja o alta complejidad, especialmente las de baja complejidad, ya que el sistema funcionará peor con ese tipo de firmas.

A continuación, se han propuesto diferentes enfoques para explotar la información relacionada con la complejidad de la firma con el objetivo final de mejorar el rendimiento de los sistemas de verificación de firma dinámica. En particular, se ha propuesto entrenar sistemas específicos por grupo de complejidad siguiendo diferentes estrategias (entrenando desde cero, probando diferentes arquitecturas del sistema, o aplicando fine tuning), pero ninguna de ellas mejoró el rendimiento del sistema de referencia TA-RNN. Por último, un enfoque basado en el entrenamiento de un sistema global con sujetos equilibrados en cuanto a su complejidad ha superado al sistema de referencia TA-RNN con una mejora relativa del 6,67 % de EER para falsificaciones cualificadas y una mejora relativa del 12,58 % de EER para falsificaciones aleatorias. Podemos concluir que el entrenamiento con un número equilibrado de sujetos en relación con la complejidad de su firma reduce el número de usuarios necesarios para lograr un rendimiento concreto, lo que hace que el entrenamiento sea más eficiente.

El análisis de los efectos de la complejidad también se ha realizado para firmas realizadas con el dedo. De esta investigación deducimos que, para sistemas entrenados exclusivamente con

firmas hechas con el dedo, el uso de un número equilibrado de sujetos en cuanto a la complejidad de su firma en el entrenamiento nos mejora la eficiencia y el rendimiento del sistema. Sin embargo, si partimos de un sistema entrenado con firmas hechas con stylus y disponemos de pocos usuarios para el entrenamiento sobre firmas realizadas con el dedo, es mejor entrenar con todas las firmas posibles.

Tras este análisis, se procedió al desarrollo de sistemas de verificación de firma estática. Se realizaron tres clases de experimentos: la evaluación del sistema de referencia sobre la base de datos DeepSignDB, el diseño de un sistema de firma estática con arquitectura siamesa y el uso de triplet loss en entrenamiento de sistemas de firma estática. De estos experimentos los mejores resultados los obtuvo un modelo que proponía una red neuronal como etapa de comparación en combinación con el extractor de características de referencia entrenado sobre DeepSignDB. Aunque sobre falsificaciones *random* si obtuvo buenos resultados (4.53 % de EER), sobre falsificaciones *skilled* este modelo no consiguió igualar al estado del arte.

Finalmente, se llevó a cabo la fusión de sistemas dinámicos y estáticos. Se realizó la combinación de las puntuaciones resultantes de cada sistema mediante una media ponderada y se estudió los puntos óptimos de fusión. Se concluyó que la fusión de sistemas dinámicos y estáticos es un enfoque útil debido a que cada sistema aporta un tipo de información diferente y la combinación de sistemas nos permite aprovechar todos los datos disponibles. El sistema resultante de esta combinación obtuvo resultados del estado del arte, alcanzando un EER de 3.61 % en falsificaciones *skilled* y un EER de 0.97 % para falsificaciones *random* considerando un escenario donde sólo tenemos acceso a una firma inscrita.

Todavía queda mucho trabajo por hacer en cuanto a la complejidad de las firmas y la fusión de sistemas. Por un lado, se propone el estudio de otro tipo de arquitecturas de aprendizaje profundo para el desarrollo de sistemas de verificación de firma dinámica, en concreto CNNs. Asimismo, el análisis de los efectos de la complejidad puede ser profundizado al estudiar casos con acceso a más de una firma inscrita. Por otro lado, queda pendiente explorar otras aproximaciones en el desarrollo de sistemas de firma estática, como puede ser el uso de la información de complejidad. Por último, se propone el estudio de la fusión de sistemas de firma dinámica y estática a nivel de características, de forma que se entrene después una capa de comparación que tenga en cuenta toda la información disponible de la firma, creando así un único sistema de verificación de firma.

Glosario de acrónimos

- **BLSTM**: *Bidirectional LSTM*
- **BGRU**: *Bidirectional GRU*
- **CNN**: *Convolutional Neural Networks*
- **DL**: *Deep Learning*
- **DTW**: *Dynamic Time Warping*
- **EER**: *Equal Error Rate*
- **FAR**: *False Acceptance Rate*
- **FRR**: *False Rejection Rate*
- **GRU**: *Gated Recurrent Unit*
- **HMM**: *Hidden Markov Models*
- **LSTM**: *Long Short-Term Memory*
- **RNN**: *Recurrent Neural Networks*
- **SVM**: *Support Vector Machines*
- **TA-RNN**: *Time Aligned Recurrent Neural Network*

Bibliografía

[1] Miguel Caruana Montes. Verificación de firma manuscrita dinámica mediante redes neuronales recurrentes. *Universidad Autónoma de Madrid*, 2019.

[2] E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez, and V. M. Patel. Exploring body shape from mmw images for person recognition. *IEEE Transactions on Information Forensics and Security*, 12(9):2078–2089, September 2017.

[3] Alejandro Acien, Aythami Morales, Ruben Vera-Rodriguez, Ivan Bartolome, and Julian Fierrez. Measuring the gender and ethnicity bias in deep models for face recognition. In *Proc. of IAPR Iberoamerican Congress on Pattern Recognition, CIARP*, pages 584–593. Springer, November 2018.

[4] Oscar Delgado-Mohatar, Julian Fierrez, Ruben Tolosana, and Ruben Vera-Rodriguez. Blockchain and biometrics: A first look into opportunities and challenges. In *Proc. International Congress on Blockchain and Applications*, June 2019.

[5] E. Gonzalez-Sosa, A. Dantcheva, R Vera-Rodriguez, JL Dugelay, F. Bremond, and J. Fierrez. Image-based gender estimation from body and face across distances. In *Proc. International Conference on Pattern Recognition, ICPR*, December 2016.

[6] E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. Exploring facial regions in unconstrained scenarios: Experience on icb-rw. *IEEE Intelligent Systems*, 2018.

[7] Pedro Tome, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia. Variability compensation using nap for unconstrained face recognition. In *Proc. 10th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS'12)*, volume 151, pages 129–139. Springer, March 2012.

[8] R. Vera-Rodriguez, M. Blazquez, A. Morales, E. Gonzalez-Sosa, J. Neves, and H. Proenca. Facegenderid: Exploiting gender information in dcnns face recognition systems. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Workshop on Bias Estimation in Face Analytics (CVPR BEFA)*, June 2019. Best Paper Runner Up Award.

[9] R. Vera-Rodriguez, J. Mason, J. Fierrez, and J. Ortega-Garcia. Analysis of spatial domain information for footstep recognition. *IET Computer Vision*, 5:380–388, 2011.

[10] Ruben Tolosana, Ruben Vera-Rodriguez, and Julian Fierrez. Biotouchpass: Handwritten passwords for touchscreen biometrics. *IEEE Transactions on Mobile Computing*, 19(7):1532–1543, 2020.

[11] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, and Aythami Morales. Biotouchpass demo: Handwritten passwords for touchscreen biometrics. In *Proc. ACM Multimedia, ACMM*, October 2019.

[12] Ruben Tolosana, Juan Carlos Ruiz-Garcia, Ruben Vera-Rodriguez, Jaime Herreros-Rodriguez, Sergio Romero-Tapiador, Aythami Morales, and Julian Fierrez. Child-computer

interaction: Recent works, new dataset, and age detection. *arXiv preprint arXiv: 2102.01405*, 2021.

[13] Ruben Tolosana, Paula Delgado-Santos, Andres Perez-Uribe, Ruben Vera-Rodriguez, Julian Fierrez, and Aythami Morales. Deepwritesyn: On-line handwriting synthesis via deep short-term representations. In *Proc. 35th AAAI Conference on Artificial Intelligence*, February 2021.

[14] Ruben Tolosana, Ruben Vera-Rodriguez, Javier Ortega-Garcia, and Julian Fierrez. Update strategies for hmm-based dynamic signature biometric systems. In *Proc. 7th IEEE Int. Workshop on Information Forensics and Security, WIFS*, November 2015.

[15] Aythami Morales, Derlin Morocho, Julian Fierrez, and Ruben Vera-Rodriguez. Signature authentication based on human intervention: Performance and complementarity with automatic systems. *IET Biometrics*, pages 1–9, June 2017.

[16] Ruben Tolosana, Ruben Vera-Rodriguez, Carlos Gonzalez-Garcia, Julian Fierrez, Santiago Rengifo, Aythami Morales, Javier Ortega-Garcia, Juan Carlos Ruiz-Garcia, Sergio Romero-Tapiador, Jiajia Jiang, Songxuan Lai, Lianwen Jin, Yecheng Zhu, Javier Galbally, Moises Diaz, Miguel Angel Ferrer, Marta Gomez-Barrero, Ilya Hodashinsky, Konstantin Sarin, Artem Slezkin, Marina Bardamova, Mikhail Svetlakov, Mohammad Saleem, Cintia Lia Szücs, Bence Kovari, Falk Pulsmeyer, Mohamad Wehbi, Dario Zanca, Sumaiya Ahmad, Sarthak Mishra, and Suraiya Jabin. Icdar 2021 competition on on-line signature verification. *arXiv preprint arXiv:2106.00739*, 2021.

[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] 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.

[19] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales, and Javier Ortega-Garcia. Do you need more data? the deepsigndb on-line handwritten signature biometric database. In *Proc. 16th IAPR Int. Conference on Document Analysis and Recognition, ICDAR*, 2019.

[20] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia. Biometric signature verification using recurrent neural networks. In *Proc. 14th IAPR Int. Conference on Document Analysis and Recognition, ICDAR*, November 2017.

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

[22] R. Tolosana, R. Vera-Rodriguez, J. Ortega-Garcia, and J. Fierrez. Preprocessing and feature selection for improved sensor interoperability in online biometric signature verification. *IEEE Access*, 3:478 – 489, May 2015.

[23] R. Guest, M. Brockly, S. Elliott, and J. Scott. An assessment of the usability of biometric signature systems using the human-biometric sensor interaction model. *International Journal of Computer Applications in Technology*, 53(4):336–347, 2016.

[24] 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.

[25] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia. Reducing the template aging effect in on-line signature biometrics. *IET Biometrics*, 8(6):422–430, June 2019.

[26] N. Houmani, A. Mayoue, S. Garcia-Salicetti, B. Dorizzi, M.I. Khalil, M.N. Moustafa, H. Abbas, D. Muramatsu, B. Yanikoglu, A. Kholmatov, M. Martinez-Diaz, J. Fierrez, J. Ortega-Garcia, J. Roure Alcobé, J. Fabregas, M. Faundez-Zanuy, J.M. Pascual-Gaspar, V. Cardeñoso-Payo, and C. Vivaracho-Pascual. BioSecure Signature Evaluation Campaign (BSEC'2009): Evaluating On-Line Signature Algorithms Depending on the Quality of Signatures. *Pattern Recognition*, 45(3):993 – 1003, 2012.

[27] R. Tolosana, R. Vera-Rodriguez, R. Guest, J. Fierrez, and J. Ortega-Garcia. Exploiting complexity in pen- and touch-based signature biometrics. *International Journal on Document Analysis and Recognition*, (23):129–141, 2020.

[28] Diaz, M., Ferrer, M.A., Impedovo, D., Malik, M.I., Pirlo, G. and Plamondon, R. A Perspective Analysis of Handwritten Signature Technology. *ACM Computing Surveys*, 51:1–39, 2019.

[29] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. BioTouchPass2: Touchscreen password biometrics using time-aligned recurrent neural networks. *IEEE Transactions on Information Forensics and Security*, 5:2616–2628, 2020.

[30] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia. Deepsign: Deep on-line signature verification. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2021.

[31] S. Otte, M. Liwicki and D. Krechel. Investigating Long Short-Term Memory Networks for Various Pattern Recognition Problems. *Machine Learning and Data Mining in Pattern Recognition*, Springer, 2014.

[32] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. Exploring Recurrent Neural Networks for On-Line Handwritten Signature Biometrics. *IEEE Access*, 6:5128–5138, 2018.

[33] S. Lai and L. Jin. Recurrent Adaptation Networks for Online Signature Verification. *IEEE Trans. on Information Forensics and Security*, 14(6):1624–1637, 2018.

[34] K. Ahrabian and B. Babaali. Usage of Autoencoders and Siamese Networks for Online Handwritten Signature Verification. *Neural Computing and Applications*, pages 1–14, 2018.

[35] Amr Hefny and Mohamed Moustafa. Online Signature Verification Using Deep Learning and Feature Representation Using Legendre Polynomial Coefficients. In *Proc. International Conference on Advanced Machine Learning Technologies and Applications*, 2019.

[36] X. Wu, A. Kimura, B.K. Iwana, S. Uchida and K. Kashino. Deep Dynamic Time Warping: End-to-End Local Representation Learning for Online Signature Verification. In *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, 2019.

[37] C. Li, X. Zhang, F. Lin, Z. Wang, J. Liu, R. Zhang and H. Wang. A Stroke-based RNN for Writer-Independent Online Signature Verification. In *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, 2019.

[38] C. Sekhar, P. Mukherjee, D.S. Guru and V. Pulabaigari. OSVNet: Convolutional Siamese Network for Writer Independent Online Signature Verification. In *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, 2019.

[39] Songxuan Lai, Lianwen Jin, Luojun Lin, Yecheng Zhu, and Huiyun Mao. SynSig2Vec: Learning Representations from Synthetic Dynamic Signatures for Real-World Verification. In *Proc. AAAI Conference on Artificial Intelligence*, 2020.

[40] Chirag Nathwani. Online Signature Verification Using Bidirectional Recurrent Neural Network. In *Proc. IEEE International Conference on Intelligent Computing and Control Systems*, 2020.

[41] Luiz G. Hafemann, Robert Sabourin, and Luiz S. Oliveira. Learning features for offline handwritten signature verification using deep convolutional neural networks. *CoRR*, abs/1705.05787, 2017.

[42] F. Vargas, M. Ferrer, C. Travieso, and J. Alonso. Off-line handwritten signature gpds-960 corpus. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, volume 2, pages 764–768, 2007.

[43] Ortega-Garcia, J. *et al.* MCYT Baseline Corpus: A Bimodal Biometric Database. *IEE Proceedings Vision, Image and Signal Processing*, 150(6):395–401, December 2003.

[44] Meenakshi K. Kalera, S. Srihari, and Aihua Xu. Offline signature verification and identification using distance statistics. *Int. J. Pattern Recognit. Artif. Intell.*, 18:1339–1360, 2004.

[45] C. Freitas, M. Morita, L. Oliveira, E. Justino, A. Yacoubi, E. Lethelier, F. Bortolozzi, and R. Sabourin. Bases de datos de cheques bancarios brasileños. *XXVI Conferencia Latinoamericana de Informatica*, 2000.

[46] Juan Hu and Youbin Chen. Offline signature verification using real adaboost classifier combination of pseudo-dynamic features. In *2013 12th International Conference on Document Analysis and Recognition*, pages 1345–1349, 2013.

[47] Yasmine Guerbai, Youcef Chibani, and Bilal Hadjadji. The effective use of the one-class svm classifier for handwritten signature verification based on writer-independent parameters. *Pattern Recognition*, 48(1):103–113, 2015.

[48] Yasmine Serdouk, Hassiba Nemmour, and Youcef Chibani. New gradient features for offline handwritten signature verification. In *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, pages 1–4, 2015.

[49] Mustafa Berkay Yılmaz and Berrin Yanıkoğlu. Score level fusion of classifiers in off-line signature verification. *Information Fusion*, 32:109–119, 2016. SI Information Fusion in Biometrics.

[50] Amir Soleimani, Babak N. Araabi, and Kazim Fouladi. Deep multitask metric learning for offline signature verification. *Pattern Recognition Letters*, 80:84–90, 2016.

[51] F. Alonso-Fernandez, M.C. Fairhurst, J. Fierrez, and J. Ortega-Garcia. Impact of Signature Legibility and Signature Type in Off-Line Signature Verification. *In Proc. IEEE Biometrics Symposium*, 2007.

[52] N. Houmani, Garcia-Salicetti, S., and B. Dorizzi. A Novel Personal Entropy Measure Confronted to Online Signature Verification Systems Performance. *In Proc. International Conference on Biometrics: Theory, Applications and System, BTAS*, pages 1–6, 2008.

[53] O. Miguel-Hurtado, R. Guest, and T. Chatzisterkotis. A new approach to automatic signature complexity assessment. In *Proc. IEEE International Carnahan Conference on Security Technology (ICCST)*, pages 1–7, 2016.

[54] A. Fischer and R. Plamondon. Signature Verification based on the Kinematic Theory of Rapid Human Movements. *IEEE Transactions on Human-Machine Systems*, 47(2):169–180, 2017.

[55] R. Vera-Rodriguez, R. Tolosana, M. Caruana, G. Manzano, C. Gonzalez-Garcia, J. Fierrez, and J. Ortega-Garcia. DeepSignCX: Signature complexity detection using recurrent neural networks. In *Proc. 15th International Conference on Document Analysis and Recognition, ICDAR*, September 2019.

[56] J. Fierrez, J. Galbally, J. Ortega-Garcia, and *et al.* BiosecurID: A Multimodal Biometric Database. *Pattern Analysis and Applications*, 13(2):235–246, 2010.

[57] Javier Galbally, Moises Diaz-Cabrera, Miguel A. Ferrer, Marta Gomez-Barrero, Aythami Morales, and Julian Fierrez. On-line signature recognition through the combination of real dynamic data and synthetically generated static data. *Pattern Recognition*, 48(9):2921–2934, 2015.

[58] 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.

[59] 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.

[60] Gustavo Manzano Ramirez. Verificación de firma manuscrita estática mediante redes neuronales convolucionales. *Universidad Autónoma de Madrid*, 2019.

[61] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. *Presentation Attacks in Signature Biometrics: Types and Introduction to Attack Detection, Handbook of Biometric Anti-Spoofing (2nd Edition)*. Springer, 2018.