



UNIVERSIDAD
DE LA REPUBLICA
URUGUAY

UNIVERSIDAD DE LA REPÚBLICA
FACULTAD DE INGENIERÍA



Fusión biométrica

Trabajo final - Introducción al Reconocimiento de Patrones

Luis Di Martino
Gabriel Lema

Tutores: Alicia Fernández, Guillermo Carbajal

MONTEVIDEO - URUGUAY
DICIEMBRE 2014

Resumen

En este trabajo se busca desarrollar un sistema completo de reconocimiento de patrones, en particular explorar las distintas técnicas de selección de características y de fusión de resultados aplicados a biometría basada en caras y huellas dactilares.

Para lograr esto se cuenta con un resultado intermedio de la comparación de caras de la base FERET, y el score final entre huellas con datos sobre la calidad de la huella de la DNIC. De modo que el objetivo primario es superar los resultados que se obtienen al utilizar los datos crudos para cada uno de los casos. Para esto primero se busca alguna forma de fusionar los resultados intermedios en la comparación de caras de forma de generar un mejor resultado y en simultáneo, se prueba cuáles son las técnicas de fusión que devuelven mejores resultados para una base de datos de cara y de huellas con datos sobre la calidad. Finalmente se utiliza dicha fusión junto con los mejorados datos de comparación entre caras para superar la performance de cada sistema biométrico por separado. Se encontró que en modo identificación la fusión logró mejorar el RR alcanzando la mejor performance posible, en tanto en modo verificación para un FAR de 10^{-4} se obtuvo una mejora de 10 puntos porcentuales.

El trabajo se divide de la siguiente forma: el Capítulo 1 realiza un acercamiento al tema y explica algunos de los términos usados en los siguientes capítulos. En el Capítulo 2 se realiza una breve descripción de los datos con los que se cuenta para hacer las pruebas de las técnicas propuestas. En el Capítulo 3 se muestran los resultados originales obtenidos con los distintos sistemas que se buscan mejorar en las etapas posteriores. El Capítulo 4 plantea la serie de pruebas realizadas en fase de entrenamiento de las estrategias de fusión propuestas. La validación de los resultados obtenidos se hace sobre una base de pruebas, los resultados obtenidos se detallan en el Capítulo 5. Finalmente en el Capítulo 6 se concluye el trabajo y se plantean líneas de trabajo a futuro.

Capítulo 1

Introducción

La biometría es una disciplina que se enmarca en las áreas de tratamiento de señales, reconocimiento de patrones y procesamiento de imágenes. Tiene como objetivo el poder caracterizar la identidad de las personas basándose en rasgos físicos o conductuales de las mismas. Estos sistemas funcionan básicamente realizando tres tareas: enrolamiento, identificación y verificación de identidades. En la etapa de enrolamiento cada muestra es preprocesada y se le realiza la extracción de características. De esta forma quedan guardados en la base de datos el vector de características con su respectiva identidad. El repositorio de datos de los individuos conocidos por el sistema se denomina comúnmente “galería” y el conjunto de datos del cual desconocemos su identidad se llama “prueba”. En el modo identificación se toma una muestra de la base de prueba, y se intenta determinar la identidad de la misma. Para esto, se comparan características extraídas de la muestra con las de la galería, y se devuelve la identidad de la comparación que obtuvo el mejor score. En el modo verificación una persona declara una identidad, y por lo tanto el sistema extrae las características, y las compara con las de su galería con esa identidad. En función de esta comparación, el sistema valida si la persona es un genuino o un impostor.

Existen varios rasgos biométricos que pueden ser utilizados, dentro de estos hay dos que son especialmente populares: las huellas dactilares y la imagen de la cara de las personas. La impresión de las huellas dactilares es muy utilizada por su gran efectividad mientras que el rostro de una persona es común por la fácil adquisición del mismo (es un rasgo que puede ser obtenido aún sin la colaboración, o conocimiento, de la persona) y es el rasgo que más comúnmente utilizamos los humanos para identificar a quienes nos rodean.

La fusión de distintos rasgos biométricos es de especial interés para organismos que trabajan con ellos para la identificación de ciudadanos. En Uruguay el mayor organismo en esta área es la Dirección Nacional de Identificación Civil (D.N.I.C) que almacena de los habitantes del Uruguay distintos rasgos biométricos (imagen de la cara, impresión de huellas dactilares, firma). Estos datos los usa al momento de validar la identidad declarada por una persona (proceso de verificación) en el momento que se emite un documento como la cédula de identidad o el pasaporte.

Capítulo 2

Bases de datos y sistemas biométricos utilizados

2.1. Bases de datos

Para evaluar las distintas técnicas propuestas en el proyecto se utilizan dos bases de datos correspondientes a distintas modalidades biométricas: reconocimiento facial y reconocimiento de huellas dactilares. Las muestras utilizadas para la evaluación del sistema de reconocimiento facial provienen de la base pública FERET [7]. Las muestras de huellas dactilares utilizadas se obtuvieron de la base privada de la DNIC. A continuación se hace una breve descripción de cada una de las bases de datos.

2.1.1. Base de caras FERET

La base de caras FERET contiene imágenes de las caras de 1010 individuos (identidades) de los cuales múltiples tomas fueron realizadas. Se provee un subconjunto de imágenes, llamado *fa*, como galería de la base de datos. El mismo contiene las imágenes que son tomadas como referencia de cada identidad para poder medir el desempeño del sistema de reconocimiento facial sobre los distintos conjuntos de prueba. En este proyecto se utiliza el subconjunto de prueba *fb* que incluye imágenes de las caras de las personas tomadas el mismo día que fueron tomadas las imágenes en la galería e incluyen diferencias de expresiones con respecto a estas. Para este proyecto se consideraron únicamente las 847 identidades que están tanto en el conjunto de prueba como la galería y, para las cuales, se incluye en la base de datos la información de las coordenadas de las posiciones de los ojos en las imágenes. Estas posiciones son utilizadas para normalizar la imagen en el funcionamiento del sistema de reconocimiento facial.

El sistema de reconocimiento facial del que se parte es Faceval [2] desarrollado principalmente por los autores de este trabajo. La extracción de características se basó en [1] donde a la imagen de entrada se le halla la imagen LBP, luego se divide la imagen LBP en parches, y a cada parche se le halla su histograma. Finalmente la comparación entre dos personas resulta en un vector de N distancias, donde N es el número de parches en los que se divide la imagen, según se muestra en la Figura 2.1.

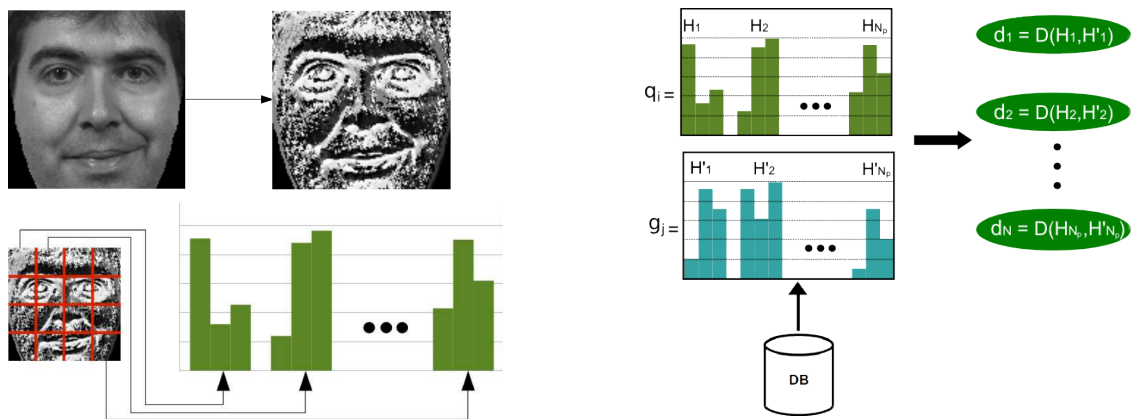


Figura 2.1: Extracción de características y comparación de vectores de características. Los datos con los que se cuenta son con las N distancias.

2.1.2. Base de huellas DNIC

La DNIC cuenta con una base de datos de huellas correspondientes a todos los individuos que en algún momento han obtenido la *Cédula de identidad (C.I.)*. Este documento, probatorio de la identidad de la persona, es obtenido por todos los ciudadanos uruguayos y extranjeros residentes en el país. Esto hace que la base de huellas con la que cuenta la DNIC sea muy grande (más de 3 millones de identidades).

Las huellas son obtenidas por el método convencional de rodado de la huella con tinta sobre papel en lo que se suelen llamar las *fichas decadactilares*. En estas fichas se adquieren las impresiones de los diez dedos de la persona al momento de tramitar la *C.I.*. Las imágenes de las huellas son posteriormente escaneadas utilizando una resolución de $500dpi$ (500 puntos por pulgada), esta resolución es la exigida comúnmente por los sistemas biométricos basados en imágenes de huellas dactilares. Cuando un individuo renueva el documento de identidad, se capturan únicamente las huellas correspondientes a los dedos pulgar e índice de las manos derecha e izquierda. Estas 4 tomas son obtenidas con un sensor electrónico (utilizando una resolución de $500dpi$) y comparadas con los registros de los dedos correspondientes de la persona que fueron obtenidos cuando se tramitó el documento por primera vez. utiliza un conjunto de prueba compuesto por 847 pares de huellas correspondientes al dedo pulgar. Cada uno de estos pares contiene la captura del mismo dedo realizada por los dos métodos descriptos anteriormente. Las huellas obtenidas de las fichas decadactilares se utilizan como conjunto de referencia (galería) y las huellas capturadas con el sensor electrónico se utiliza para medir el desempeño (conjunto de prueba).

Cuando se comparan huellas, además de contar con el resultado de la comparación (score o distancia), se tienen indicadores de la calidad de cada una de las muestras utilizadas, tanto de la muestra en el conjunto de prueba como de la muestra en la galería contra la cual la primera se compara. Comúnmente, los indicadores de calidad no son utilizados en la clasificación. Son utilizados generalmente en la etapa de preprocesamiento de sistemas de reconocimiento de huellas dactilares para descartar o solicitar una retoma de la muestra cuando la misma no cumple con los requerimientos mínimos de calidad establecidos. En este proyecto se considera la posibilidad de que la información provista de la calidad de las huellas puede ser útil al momento de clasificar una comparación entre dos muestras cuando se está realizando la verificación de identidades.

2.1.3. Abordaje al problema

Para este proyecto se divide tanto la base de huellas DNIC como la base de caras FERET, en dos conjuntos; TRAIN y TEST. La idea es realizar el aprendizaje de parámetros, selección del mejor método, o cualquier tipo de ajuste utilizando el conjunto TRAIN. El conjunto TEST se utiliza únicamente una vez que se haya finalizado la etapa de aprendizaje y ajuste, para validar lo hecho. Vale la pena aclarar que tanto el conjunto TRAIN como TEST están compuestos por una galería y un conjunto de prueba, donde las identidades del conjunto de prueba coinciden con las de la galería.

Capítulo 3

Desempeño de los sistemas originales

Para poder evaluar las distintas técnicas de fusión propuestas, es necesario conocer primero el desempeño obtenido por cada sistema biométrico utilizado de forma individual. A continuación se muestran los resultados obtenidos cuando cada sistema es utilizado de forma independiente en las tareas de verificación e identificación. Los resultados obtenidos por cada sistema se muestran juntos de forma de poder comparar el desempeño que obtiene cada sistema.

3.1. Modo identificación

A continuación se muestra el desempeño obtenido cuando se utiliza cada sistema en el modo identificación de forma individual. Para esta evaluación, el sistema de reconocimiento facial es utilizado asignando a todos los parches en las imágenes de las caras el mismo peso en la clasificación. Esto es equivalente a pensar que no existen zonas de la cara que son más o menos discriminantes al momento de comparar dos muestras. Con respecto al sistema de reconocimiento de huellas dactilares, se utilizan únicamente los scores obtenidos por el sistema de clasificación, no se consideran en esta prueba los valores de calidad asociados a cada huella.

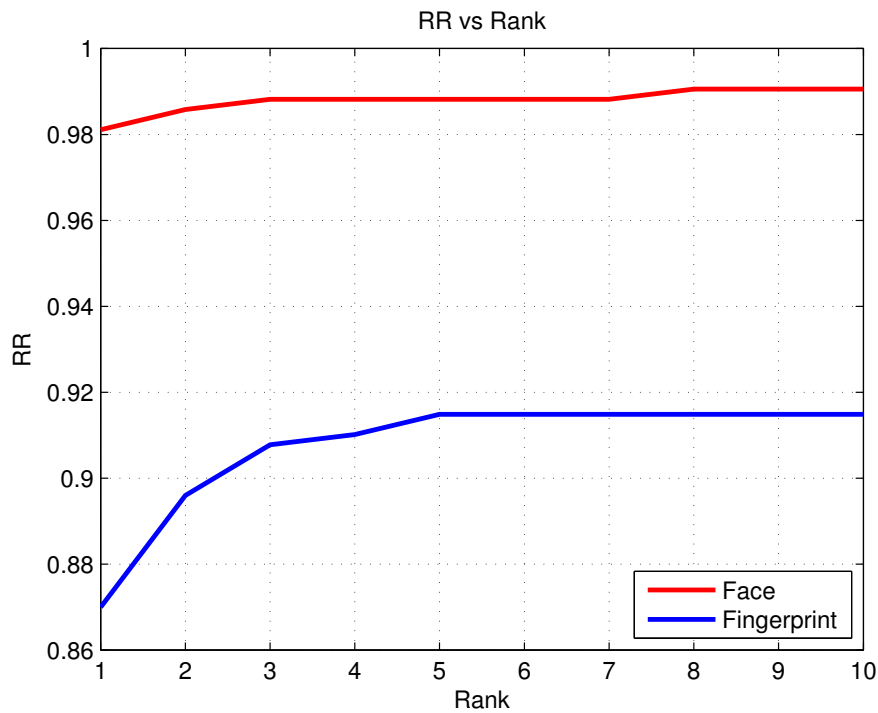


Figura 3.1: RR vs. Rank, modo identificación, conjunto TEST

Como se puede ver en la Figura 3.1, el sistema de reconocimiento facial logra un mejor desempeño en la tarea de identificación que el sistema de reconocimiento de huellas dactilares.

3.2. Modo verificación

A continuación se muestra el desempeño obtenido cuando se utilizan los sistemas de forma individual para realizar la verificación de identidades. Una vez más, se considera en este caso que todos los parches de la cara tienen igual peso al momento de realizar la clasificación y no se consideran los valores de calidad de las huellas comparadas.

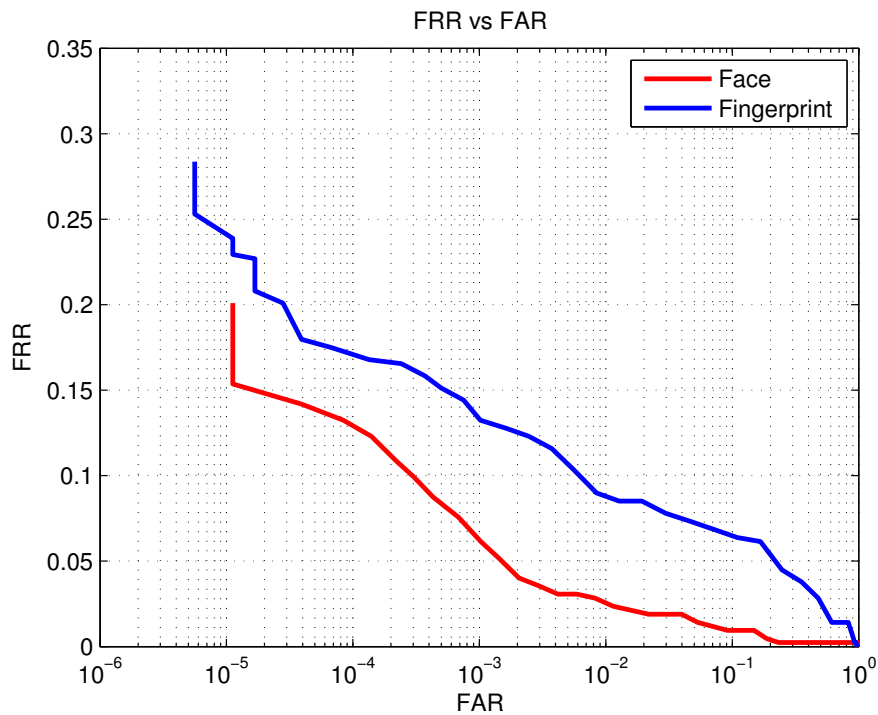


Figura 3.2: FRR vs. FAR, modo verificación, conjunto TEST

El desempeño de un sistema biométrico que trabaja en el modo de verificación se basa en que tan buena separación logre el mismo entre las clases “impostores” y “genuinos” en las cuales puede ser clasificada la comparación de dos muestras biométricas. Siendo que se utilizan bases de datos de caras y huellas correspondientes a $n = 423$ identidades, se tienen n muestras de la clase genuinos y $n(n - 1)$ de la clase impostores. A continuación se muestran la distribución de estas clases para ambos sistemas utilizados, los histogramas siguientes fueron obtenidos utilizando 100 muestras de la clase genuinos y 1000 de la clase impostores, los resultados mostrados corresponden a considerar las salidas de los sistemas como scores (un valor grande como resultado de una comparación indica pertenencia de la comparación a la clase genuinos) normalizados en el intervalo $[0, 1]$.

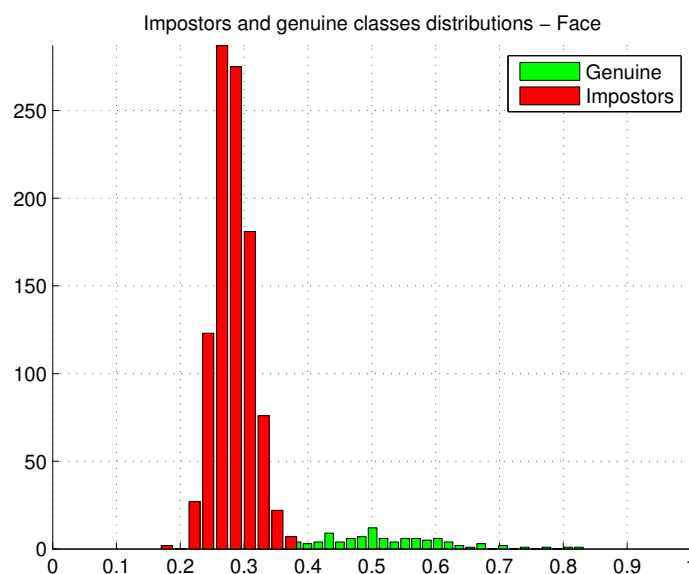


Figura 3.3: Distribución de clases, sistema de reconocimiento facial, conjunto TEST

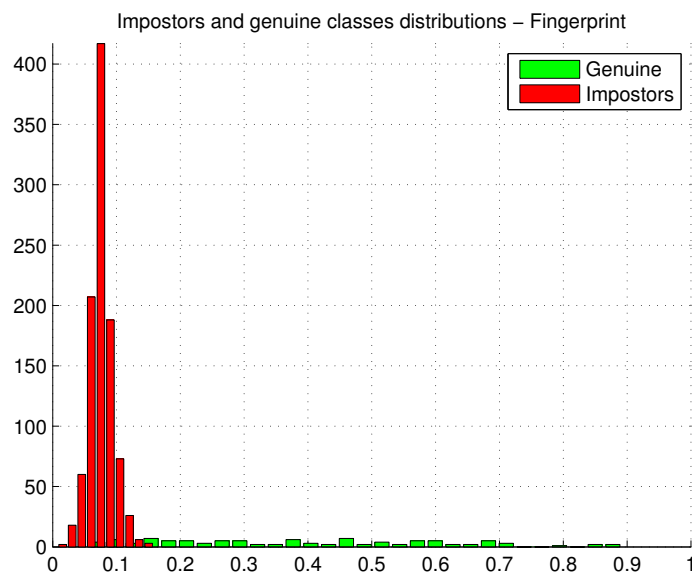


Figura 3.4: Distribución de clases, sistema de reconocimiento de huellas dactilares, conjunto TEST

Como se puede ver, a partir de las imágenes anteriores, el sistema de reconocimiento facial también presenta un mejor desempeño cuando se utilizan los sistemas en el modo de verificación, logrando una menor tasa de falsos rechazos (FRR) para un mismo valor de la tasa de falsas aceptaciones (FAR).

Capítulo 4

Entrenamiento - Experimentos y resultados

El objetivo de este capítulo es encontrar la mejor selección de atributos que utiliza cada sistema por separado y su mejor forma de fusión. Posteriormente, se busca obtener la mejor fusión de los resultados finales de cada sistema individual obteniendo como resultado un sistema biométrico multi-modal basado en caras y huellas. Los experimentos se realizan utilizando únicamente la base TRAIN, asumiendo que la base TEST es desconocida en esta etapa.

Los métodos de fusión basados en scores utilizan los resultados de las comparaciones obtenidas de cada sistema y los fusionan de acuerdo a distintas reglas de combinación. Para utilizar estos métodos se pueden utilizar tanto scores como distancias entre muestras, dependiendo de las salidas de los sistemas con los que se trabaja. Un sistema que utiliza scores en la etapa de clasificación asigna un valor alto en la salida para indicar que las dos muestras comparadas corresponden a la misma clase. Por el contrario, cuando distancias son utilizadas, un alto valor en la salida indica que la comparación realizada es entre muestras de distintas clases.

A diferencia de los métodos de fusión basados en *rankings* o basados en las decisiones de cada sistema, los métodos de combinación de *scores* no tienen en cuenta los resultados cualitativos (asignación de una etiqueta) de cada sistema. Simplemente, utilizan los resultados numéricos de las comparaciones para obtener un nuevo score (o distancia) asociada a la comparación de dos identidades. Finalmente, la estimación del desempeño de la fusión se realiza sobre los scores obtenidos. Como en este proyecto se cuenta con las salidas de los sistemas biométricos utilizados como scores o distancias entre muestras, resulta natural abordar el problema de fusión utilizando técnicas que hagan uso de esta información y se descarta el uso de otras técnicas de fusión (basadas en múltiples instancias del rasgo biométrico, múltiples sensores o múltiples características de un mismo rasgo biométrico).

Un sistema biométrico puede buscar resolver tanto el problema de identificación como de verificación. Si bien el resultado final depende del modo utilizado, en ambos casos se puede modelar al problema como uno de clasificación de una comparación entre muestras biométricas en dos clases: “*impostores*” y “*genuinos*”. La salida del sistema puede ser una etiqueta (en el caso de que se pretenda realizar la verificación de identidades) y/o un valor de score (o distancia) de la comparación a cada una de las clases. Este valor, que indica la pertenencia de la comparación a cada una de las clases, puede ser utilizada posteriormente para realizar la identificación. La fusión de los sistemas puede plantearse entonces dos niveles distintos:

- Fusión de scores de cada sistema individual.
- Fusión de los scores finales de cada sistema conformando un sistema multi-modal.

4.1. Fusión de scores - Sistemas individuales

Los sistemas utilizados dan como resultado de una comparación más que un simple valor de distancia (o score) entre las muestras de entrada. En particular, cuando se utiliza el sistema de reconocimiento facial se obtienen los resultados de la comparación de los 81 parches entre las caras y, cuando se utiliza el sistema de reconocimiento de huellas dactilares, se obtienen indicadores de calidad de las huellas comparadas. Esta información puede ser utilizada considerando los resultados que obtiene cada sistema de forma individual como un vector de características de la comparación entre las muestras que debe ser clasificado en la clase impostores o genuinos.

Para las siguientes pruebas se utiliza la base de entrenamiento TRAIN conteniendo 424 muestras de la clase genuinos, y 424×423 de la clase impostores. A continuación se presenta el esquema propuesto para realizar la fusión a nivel de scores.

4.1.1. Ponderación de los parches en la cara

Se aborda el problema de ponderación de parches como uno de selección de características donde se busca aquellas que son las más relevantes al momento de clasificar la comparación entre dos caras en las clases impostores y genuinos. Para hacer las pruebas se utiliza, debido al gran desbalance de clases, una muestra impostora por cada muestra genuina; para cada persona en la base de prueba existen 423 personas en la galería que son impostoras, y se elige una al azar. Se proponen los siguientes métodos usuales de selección de características y extracción de características:

- Wrapper: evalúan la utilidad de distintos subconjuntos de características, en base al desempeño que se logra con un determinado clasificador. Se trabajó con un clasificador Random Forest y una búsqueda Best First
- Filter: evalúan la relevancia de cada característica de forma individual. A partir de un subconjunto de características se incluye o se quita una característica según una función de mérito. En particular se utilizó el test de relevancia estadística Chi-cuadrado, información mútua, y ganancia de información.
- AdaBoost: este método permite crear un clasificador fuerte, a partir de varios clasificadores débiles. Esto lo hace seleccionando un grupo de características al azar, hallando el error de clasificación, y luego repitiendo el proceso pero dándole más probabilidad de ser seleccionadas a aquellas características que fueron mal clasificadas, de forma que en el próximo paso la clasificación de mejor resultado. Si bien AdaBoost tiene como objetivo la creación de un clasificador fuerte, se pueden usar los resultados intermedios para ver cuáles características son aquellas que mejor discriminan las clases. En particular se utilizó como clasificador débil un árbol de decisión de un sólo nodo.

Una vez que se encuentran las características que mejor representan a cada clase, se vuelve nuevamente al problema original donde se tiene para cada persona la distancia entre parches, y se busca la forma de fusionar esas distancias. Para esto se utilizan los siguientes métodos:

- Usual: suma de distancias sin ponderación.
- LDA: busca el vector de proyección que maximiza la varianza interclase y minimiza la varianza intraclase. El método de fusión está basado en [9].

- SVM lineal: busca encontrar el hiperplano que mejor separa las muestras minimizando una función de costo. Una vez que se encuentra el hiperplano se utiliza el vector normal al hiperplano, como vector de proyección.

La fusión se realiza de la siguiente forma: dada una persona de la base de prueba q_k y una persona de la galería g_j , la distancia entre ellas es dada por

$$D(q_k, g_j) = \sum_i^N \omega_i d_i$$

donde

- N es el número de parches utilizados
- d_i son las distancias entre parches de esas personas
- ω_i es el vector que indica el peso de esa distancia entre parches. Para el caso LDA o SVM viene dado por el vector de proyección, para el caso usual es un vector de largo N donde cada entrada es 1.

Debido a que realizar cada una de las combinaciones de selección de características y fusión llevaría mucho tiempo, se decidió que se haría la búsqueda de características que de el mejor resultado utilizando la fusión usual. Esta decisión se fundó en el hecho de que la fusión usual es la más rápida debido a que no requiere un entrenamiento, y además es la más sencilla de interpretar. Por este motivo se omitió la selección de características por métodos Wrapper, debido a que no el clasificador no era entrenable. Luego para esa selección de características se probará las distintas formas de fusión.

Resultados - Fusión a nivel de características

A continuación se presentan los resultados que se obtuvieron para cada método de selección o extracción de características.

Wrapper: Se utilizó la implementación de WEKA, utilizando como clasificador Random Forests. La Figura 4.1 muestra los parches seleccionados, donde los parches con mayor intensidad son los que quedaron en primer lugar.

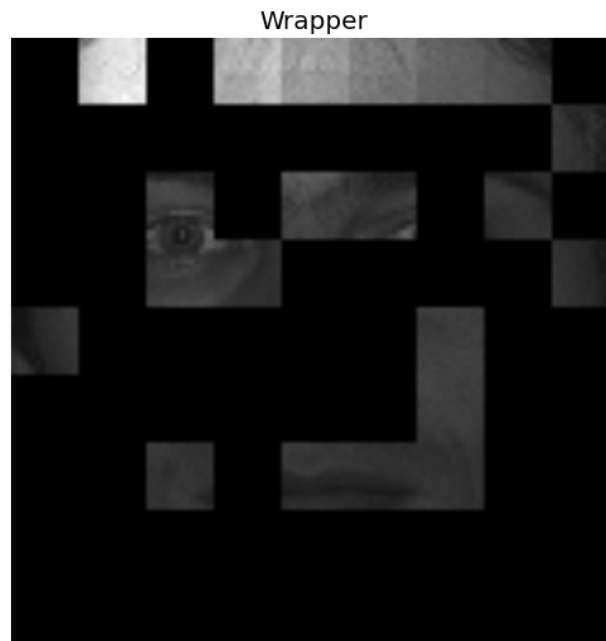


Figura 4.1: Parches extraídos utilizando wrapper

Filter: Los métodos usados, implementados en WEKA, son:

- Ganancia de Información:

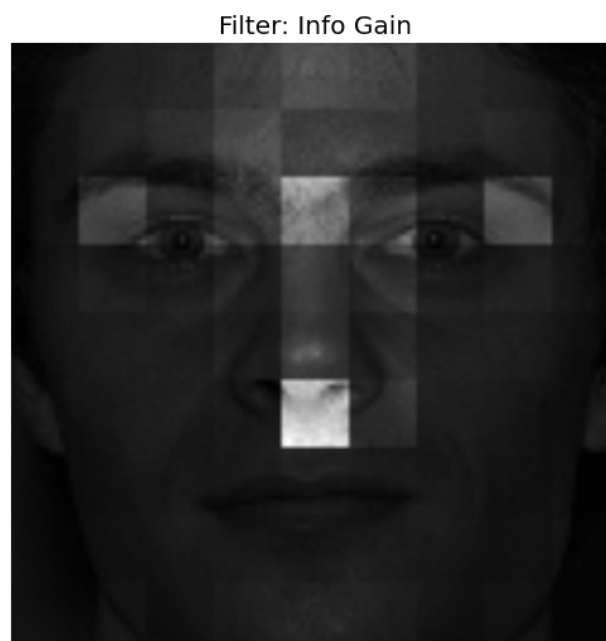


Figura 4.2: Parches extraídos utilizando ganancia de información wrapper

- Razón de ganancia

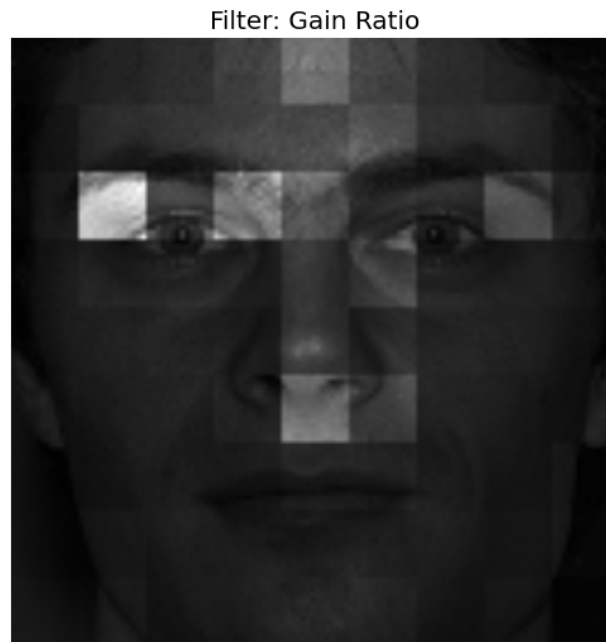


Figura 4.3: Parches extraídos utilizando razón de ganancia

- Chi-Cuadrado



Figura 4.4: Parches extraídos utilizando Chi-cuadrado

Adaboost: La Figura 4.5 muestra los parches seleccionados utilizando AdaBoost.

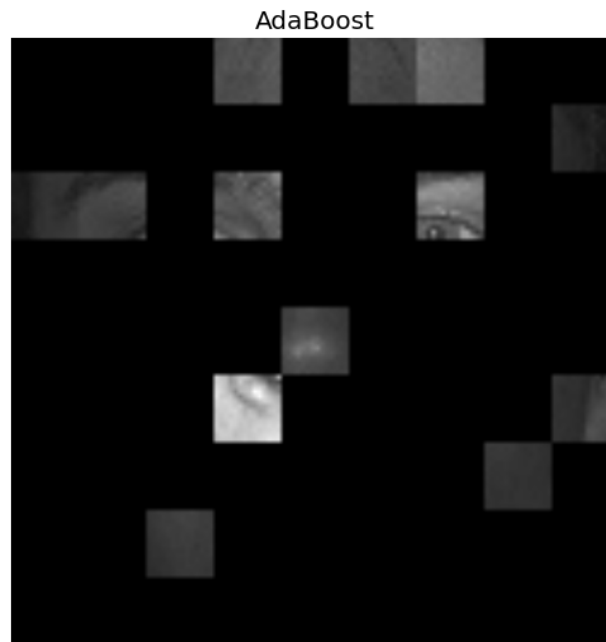


Figura 4.5: Parches extraídos utilizando AdaBoost

En primer lugar se puede concluir que existe una cierta concordancia entre los métodos Filter de cuáles son los parches más importantes. En particular podemos observar la existencia de simetría axial vertical. Por un lado es alentador en el sentido que encontró que si uno era importante, su simétrico también lo era. Por otro lado, uno esperaría que si ya se encontró uno, el otro simétrico no aportara información, es decir que no hubiera ganancia de información, o que no fuera estadísticamente relevante. Los resultados obtenidos con AdaBoost son interesantes, ya que encontró en la tercera fila un área que casi cubre todo el ojo izquierdo, y el parche faltante lo encontró en su simétrico. No resulta sencilla la interpretación de por qué los parches que aparecen en la primera fila de la imagen en el caso Wrapper podrían discriminar mejor que otros como los que se encuentran en la cara misma.

Para la evaluación de cada método se propuso realizar una división de la base de datos y galería del conjunto TRAIN en diez partes. Esto se hizo para eliminar la posibilidad de haber seleccionado características que afecten a un grupo de muestras, dado que en la etapa de selección se hizo utilizando todo el conjunto de muestras. El resultado que se toma finalmente es el promedio de las diez curvas. La Figura 4.6 muestra la curva de RR vs Rank para cada uno de los métodos. Cabe aclarar que el criterio de selección de parches se basó en heurísticas o se hizo con clasificadores que no son exactamente los que se usan en la identificación. Por lo tanto era esperable que pudieran no mejorar los resultados con respecto a la selección de todos los parches encontrados. De hecho, varios trabajos que se encuentran en el estado del arte del reconocimiento facial apuntan a realizar la extracción de características utilizando un muestreo denso de parches en la cara [3],[8],[10].

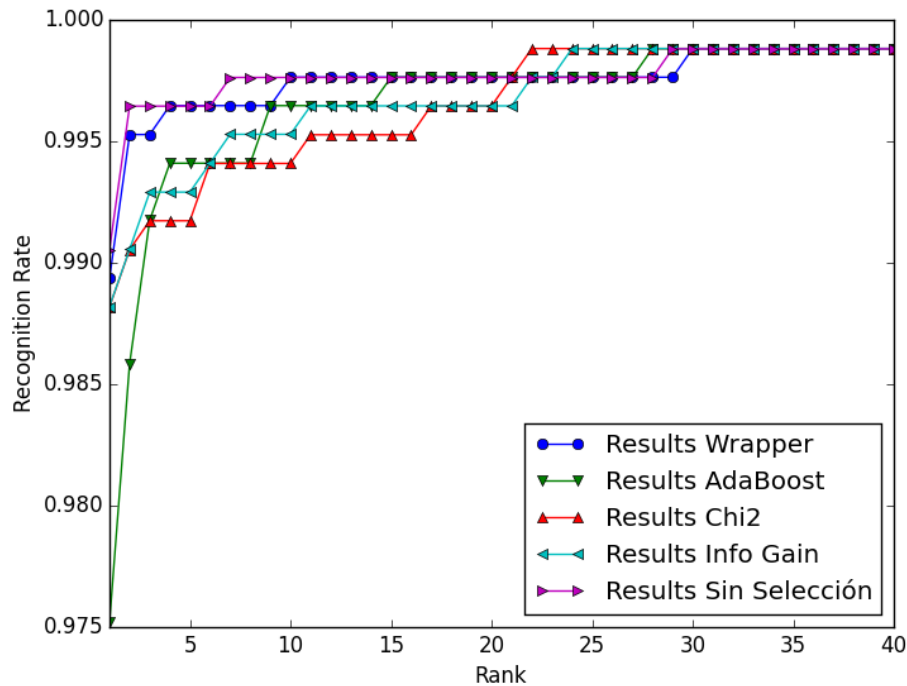


Figura 4.6: Curva RR vs Rank para cada método de selección de características

En la segunda etapa se pretende evaluar cuál de los métodos de fusión devuelve el mejor resultado. Para esto, debido a que se requiere una etapa de aprendizaje del modelo, se utiliza un esquema de *10-fold cross validation*. Para la búsqueda del mejor parámetro C del SVM lineal se realizó una búsqueda general y luego más específica usando un esquema de *5-fold cross validation* (dentro del subconjunto de muestras). Finalmente se obtuvo que el mejor valor de C es $5e - 5$. Los resultados se computaron realizando 10 corridas de cada una de esas *10-fold cross validation* y tomando el promedio.

La Figura 4.7 muestra los resultados obtenidos utilizando los distintos métodos de fusión.

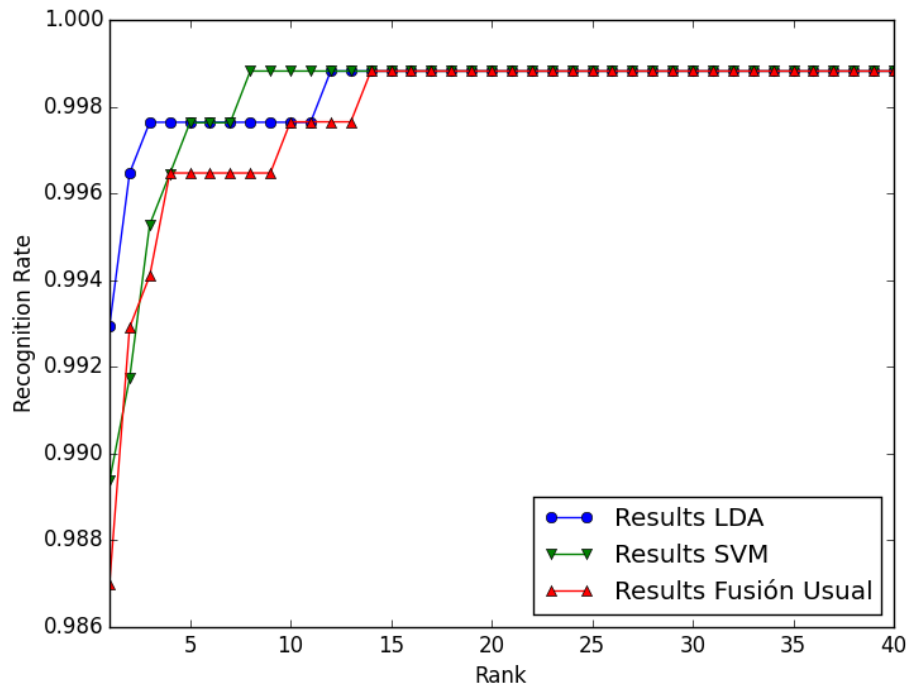


Figura 4.7: Curva RR vs Rank para cada método de selección de características

Tanto la fusión utilizando LDA, como SVM superan a la fusión usual, lo cual es un resultado alentador. A su vez se observa que SVM y LDA tienen un resultado muy similar y no resulta claro cuál es mejor, ya que en el caso de LDA, la curva está por arriba en los primeros lugares, pero luego es superado por SVM. Sin embargo, debido a que SVM para las primeras posiciones se encuentra en algunas posiciones por debajo que la fusión usual, se concluye que LDA es el mejor.

Finalmente se computa utilizando el vector promedio de cada entrenamiento de LDA, el ranking sobre toda la base de datos de TEST.

La Figura 4.8 muestra los resultados obtenidos cuando se utiliza la base de datos TEST.

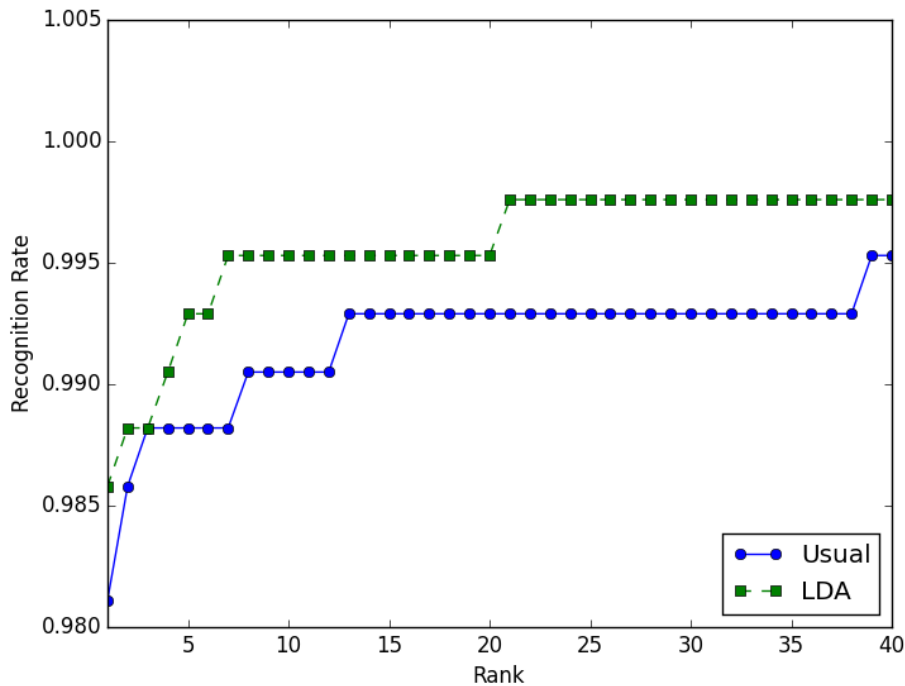


Figura 4.8: Curva RR vs Rank sobre la base de datos TEST.

Se observa que los resultados utilizando LDA son mejores que los que se obtienen utilizando fusión usual. De esta forma se valida el abordaje realizado y se concluye la etapa de encontrar el mejor sistema de reconocimiento facial en modo identificación, para los datos dados.

4.1.2. Incorporación de información de calidad de las huellas - redes neuronales

Para incorporar la información de calidad de las huellas en la etapa de clasificación, se genera por cada comparación, un vector de características que incluye: el resultado de la comparación entre las huellas (score), la calidad de la huella cuya identidad se está validando y la calidad de la huella en la galería correspondiente a la identidad declarada. Este vector de 3 características puede ser, posteriormente, clasificado en las clases “impostores” o “genuinos”. Se entrenó una red neuronal utilizando vectores obtenidos del subconjunto de entrenamiento de la base de huellas dactilares y se probó el desempeño de la misma en el subconjunto de evaluación.

A continuación se muestran los resultados obtenidos para distintas configuraciones de la red neuronal. Las arquitecturas probadas tienen la forma $I - H - O$ donde el primer término (I) corresponde a la cantidad de neuronas en la capa de entrada de la red, la última (O) a las neuronas en la capa de salida y los términos en el medio (H) representan las neuronas utilizadas en las capas ocultas (internas) de la red. Los valores FAR_{NN} y FRR_{NN} corresponden a la tasa de falsas aceptaciones y falsos rechazos obtenidos con la red neuronal, el valor FRR corresponde a la tasa de falsos rechazos que se obtiene por el sistema original, que no tiene en cuenta la calidad de las huellas, para el valor de FAR_{NN} obtenido.

Configuración red neuronal	FAR_{NN}	FRR_{NN}	FRR
1-2-1	2.24 e-5	0.20	0.20
1-2-2-1	2.8 e-5	0.18	0.20
1-2-2-2-1	3.92 e-5	0.18	0.18
1-3-1	3.92 e-5	0.18	0.18
1-4-1	3.92 e-5	0.18	0.18

La Figura 4.14 muestra la relación FRR vs FAR obtenida cuando no se considera la calidad de las huellas en la clasificación realizada por el sistema funcionando en el modo verificación sobre la base TRAIN.

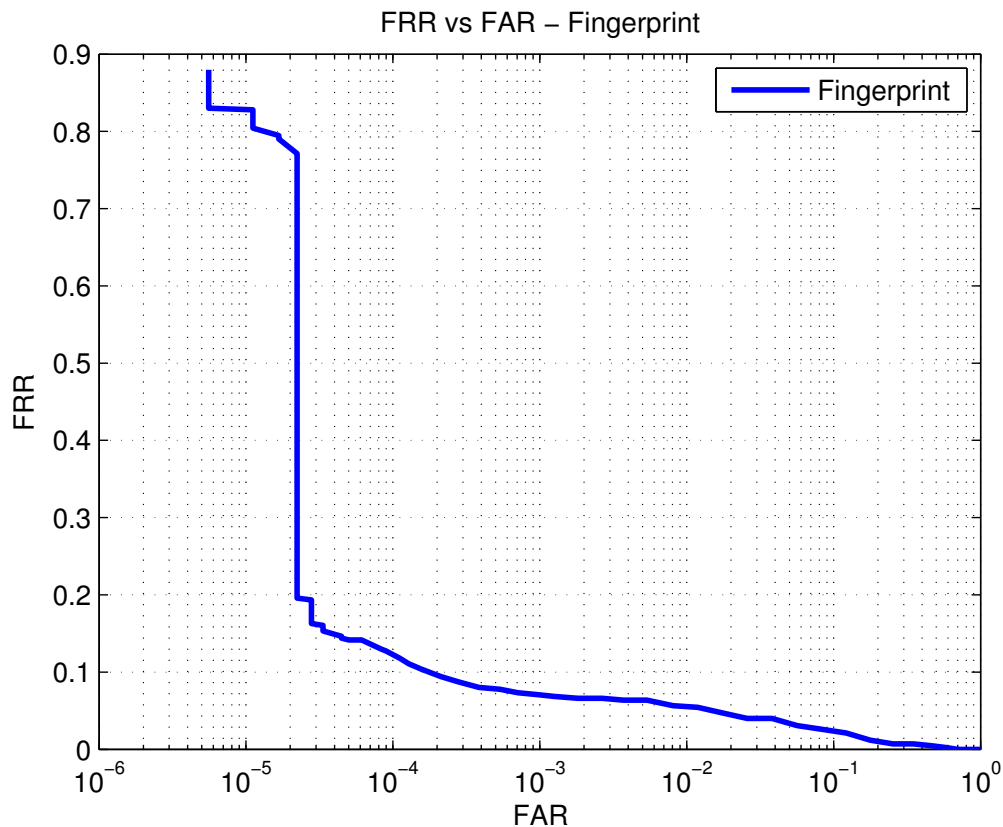


Figura 4.9: FRR vs. FAR, sistema de reconocimiento de huellas sobre la base TRAIN

Como se puede ver en la Tabla 4.1.2, los resultados obtenidos cuando se utiliza una red neuronal para incluir los indicadores de calidad en la clasificación son muy similares a los que se obtiene cuando esta información no es tenida en cuenta. Únicamente cuando se utiliza la arquitectura 1-2-2-2-1 de la red neuronal se puede ver una mejora, un menor valor de FRR para el mismo valor de FAR , aunque esta diferencia no es muy significativa. Dado que el desempeño del sistema no presenta cambios significativos cuando se incluye la información dada por la calidad de las huellas, se decide no utilizar la misma al momento de hacer la fusión.

4.2. Fusión de scores - sistema multimodal

Los métodos de fusión basados en scores utilizan los resultados de las comparaciones obtenidas de cada sistema y los fusionan de acuerdo a distintas reglas de combinación. Para utilizar estos métodos se pueden utilizar tanto scores como distancias entre muestras, dependiendo de las salidas

de los sistemas con los que se trabaja. El uso de unos u otros no cambia el funcionamiento de las reglas, simplemente es necesario que las dos entradas a la estrategia de fusión sean compatibles. En caso de que los sistemas fusionados utilicen distintos criterios, simplemente se invierten los resultados de uno de ellos para que las entradas sean compatibles.

Las reglas de combinación utilizadas no son específicas del área de la biometría. Las mismas son presentadas en [5] para la fusión de varios resultados de distintos sistemas de reconocimiento de patrones. Estas se basan en la combinación de los resultados obtenidos de L sistemas que como salida brindan la probabilidad a posteriori de la pertenencia de una muestra x a cada una de las clases en las que puede clasificar la muestra. En [6] se hace una revisión de estas estrategias para la fusión de resultados de sistemas biométricos. Si se tienen C clases en las que se puede clasificar el resultado y L sistemas que realizan la clasificación, se obtiene para cada muestra x una matriz de distancias $D(x)$. Cada elemento $d_{i,j}(x)$ de esta matriz corresponde a la distancia dada por el sistema i con respecto a la clase j . Las reglas de fusión quedan definidas de la siguiente manera.

4.2.1. Reglas de combinación de scores

- Suma: $\tilde{d}_j(x) = \sum_{i=1}^L d_{i,j}$ con $j = 1, \dots, C$
- Producto: $\tilde{d}_j(x) = \prod_{i=1}^L d_{i,j}$ con $j = 1, \dots, C$
- Min: $\tilde{d}_j(x) = \min_L (d_{1,j}, d_{2,j}, \dots, d_{L,j})$
- Max: $\tilde{d}_j(x) = \max_L (d_{1,j}, d_{2,j}, \dots, d_{L,j})$

Donde $\tilde{d}_j(x)$ representa la distancia resultante de la fusión de la muestra x a la clase ω_j . Como fue comentado anteriormente, el ajuste de estas estrategias al uso de scores entre muestras es inmediato, considerando simplemente la distancia entre dos representantes como el inverso del score obtenido entre los mismos. Si se utilizan scores en la comparación, las reglas *Min* y *Max* consideran el peor y mejor caso respectivamente de entre los sistemas fusionados. La regla del producto se basa en el hecho de que los scores son proporcionales a la probabilidad a posteriori de que una comparación entre dos muestra pertenezca a una de las posibles clases. Por tanto, hacer el producto surge como la probabilidad condicional de que ambos sistemas clasifiquen la comparación en una de las clases, asumiendo que las clasificaciones realizadas por los distintos sistemas son eventos independientes. El producto de scores puede dar malos resultados en el caso que uno de los sistemas da como salida de la comparación un score demasiado bajo (valor muy cercano a cero), ya que el mismo anula el score provisto por el otro sistema. La regla de la suma provee una alternativa que puede considerar ambos resultados en estos casos.

4.2.2. Normalización de los scores

Para poder aplicar las reglas anteriores es necesario que los datos de los sistemas a ser fusionados estén contenidos en un mismo rango numérico, por ejemplo $[0, 1]$. Esto hace necesaria la aplicación de una etapa de normalización de los resultados. En este proyecto se utilizó para la normalización de los scores la función *QLS* introducida en [4] para la normalización de resultados de sistemas biométricos. Esta función pertenece a la familia de funciones sigmoideas, aplicando la siguiente ecuación a un valor x de score entre muestras.

$$QLS(x) = \frac{ab + 1}{1 - b} \frac{1 - b^{\frac{x}{x_{max}}}}{ab^{\frac{x}{x_{max}}} + 1} \tag{4.1}$$

Donde x_{max} representa el máximo valor que puede alcanzar la comparación y a y b corresponden a los siguientes valores.

$$a = 2 + \sqrt{3} \quad y \quad b = 7 - 4\sqrt{3} \tag{4.2}$$

4.2.3. Resultados

Resultados - modo verificación

La Figura 4.10 muestra los resultados obtenidos con las distintas reglas de fusión de scores junto al desempeño que cada sistema obtiene por separado cuando funciona en el modo verificación. Los resultados se obtienen utilizando la base TRAIN.

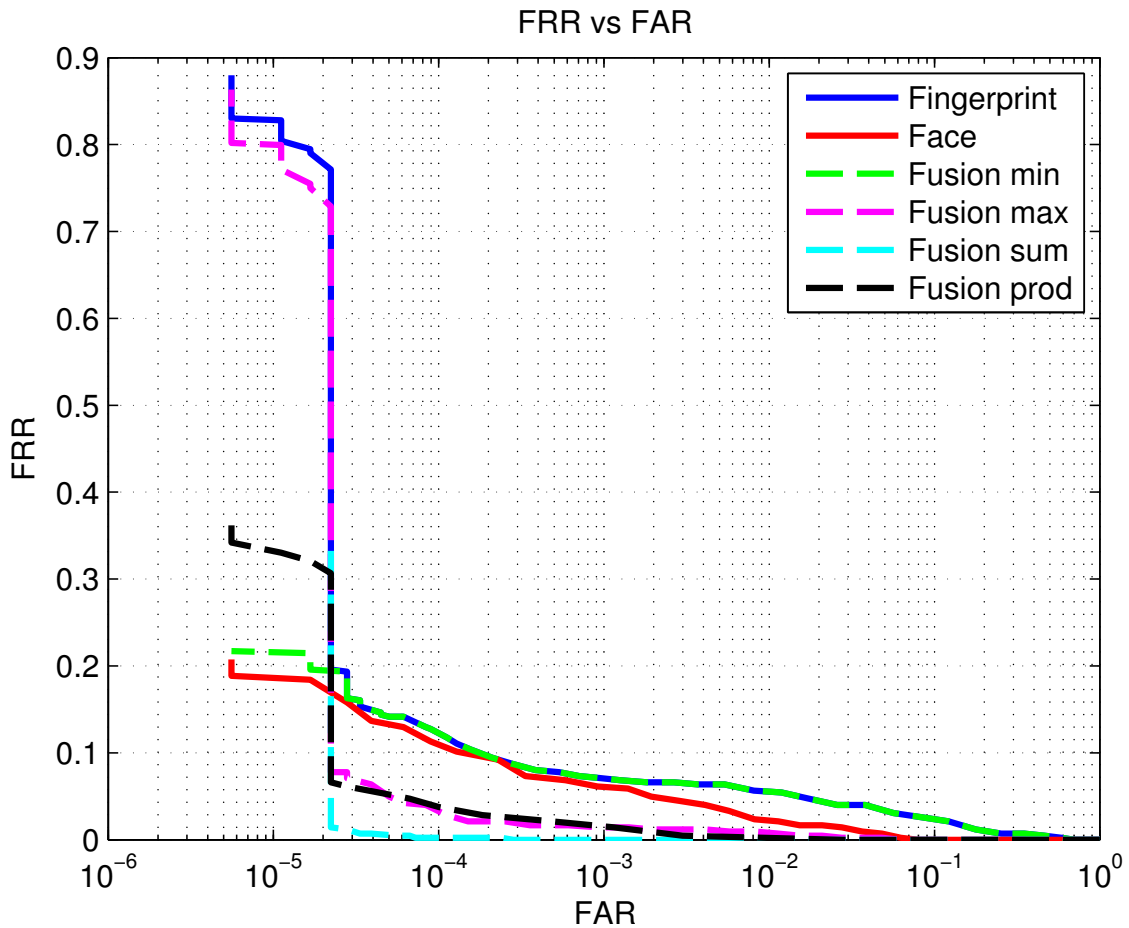


Figura 4.10: FRR vs. FAR para las distintas estrategias de fusión de scores sobre la base TRAIN.

En todas las estrategias de combinación logran un mejor resultado que el obtenido con cada sistema biométrico funcionando de forma independiente. De las reglas propuestas, la que produce el mejor resultado es la que realiza la suma de las distancias resultado de utilizar cada sistema de forma individual.

Las Figuras 4.11, 4.12 y 4.13 muestran la distribución de las clases impostores y genuinos obtenidas originalmente por los dos sistemas utilizados de forma individual sobre el conjunto de

entrenamiento y la distribución resultante de hacer la fusión utilizando la regla de suma, respectivamente.

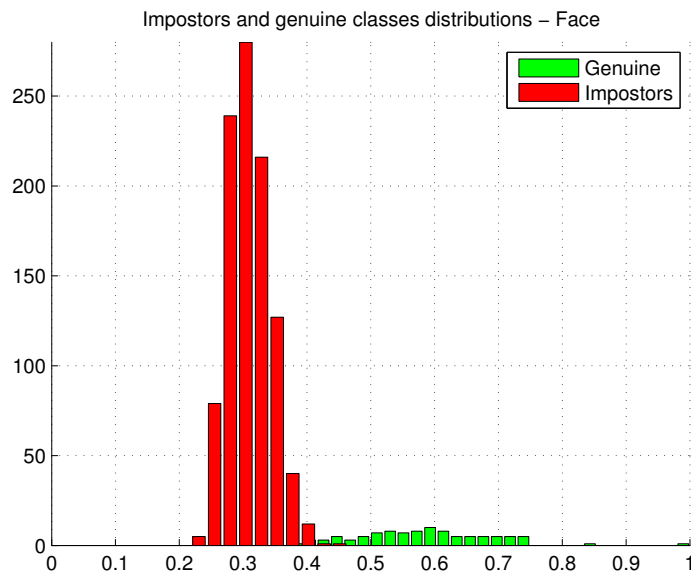


Figura 4.11: Distribución de clases, sistema de reconocimiento facial, conjunto de entrenamiento

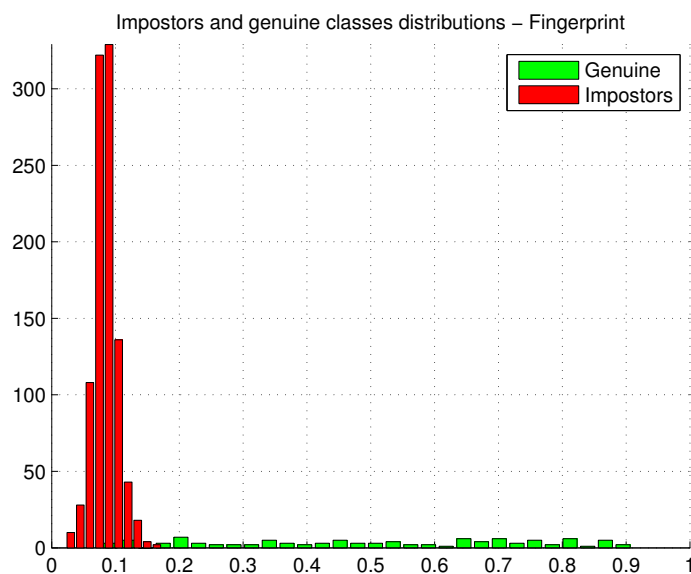


Figura 4.12: Distribución de clases, sistema de reconocimiento de huellas dactilares, conjunto de entrenamiento

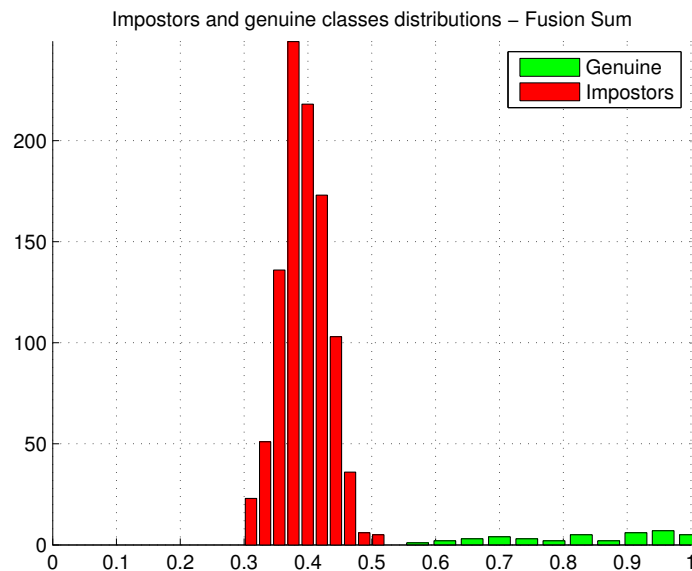


Figura 4.13: Distribución de clases, fusión por suma, conjunto de entrenamiento

Al utilizar la estrategia de suma de scores para hacer la fusión se logra una mejor separación de los representantes de las clases impostores y genuinos. Esto justifica la diferencia de desempeño con respecto al funcionamiento de los sistemas originales que se puede ver en la Figura 4.10.

Resultados - modo identificación

La Figura 4.14 muestra los resultados obtenidos con las distintas reglas de fusión de scores junto al desempeño que cada sistema obtiene por separado cuando funciona en el modo identificación. Los resultados se obtienen utilizando la base TRAIN.

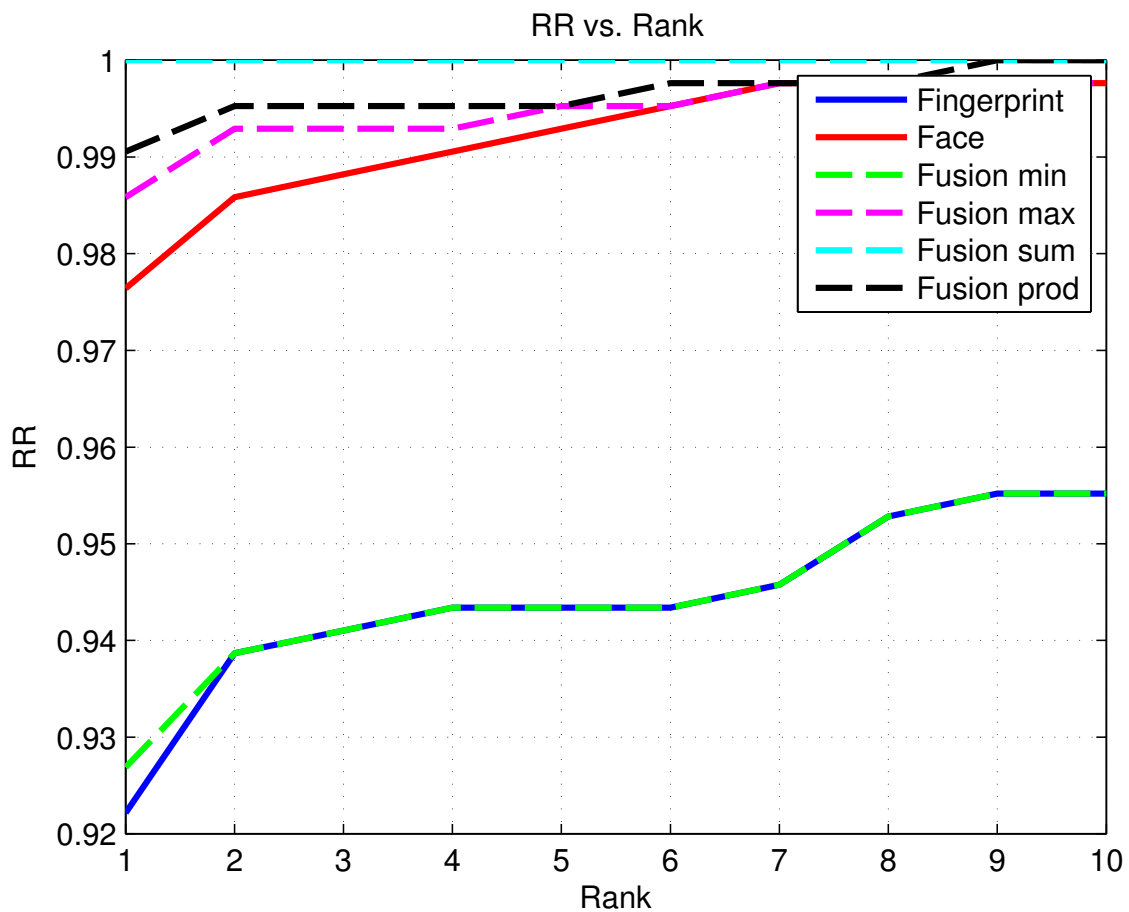


Figura 4.14: RR vs. rank para las distintas estrategias de fusión de scores sobre la base TRAIN.

El mejor desempeño se obtiene, una vez más, cuando se utiliza la suma de scores para realizar la clasificación. La regla de fusión de suma logra buenos resultados debido a que únicamente puede obtener malos resultados cuando ambos sistemas considerados cometen los mismos errores al identificar o verificar una identidad. Dado que son sistemas independientes, este evento es muy poco probable por lo tanto se logran mejorar considerablemente los resultados para ambos modos al utilizar esta estrategia.

Capítulo 5

Validación- Experimentos y resultados

En el Capítulo 4 se evaluaron, por un lado, distintas técnicas de fusión de los resultados que obtiene cada sistema biométrico por separado buscando mejorar los desempeños que cada sistema logra de forma individual. Posteriormente, se buscó la mejor forma de fusionar los resultados de ambos sistemas para conformar un único sistema biométrico multi-modal que utilizara ambos rasgos biométricos de acuerdo al objetivo planteado al inicio del proyecto. En este capítulo se validan las técnicas de fusión elegidas evaluando el desempeño de las mismas en la base de TEST. Se utilizan para las pruebas los siguientes resultados obtenidos en el Capítulo 4:

- **Pesos LDA:** se utilizan, como primer cambio, los pesos hallados anteriormente utilizando *LDA* para la combinación a través de una suma ponderada de los resultados que obtiene el sistema de reconocimiento facial utilizando los parches de la cara de forma individual.
- **Suma de distancias de ambos sistemas:** se combinan los resultados obtenidos por el sistema de reconocimiento facial (incluyendo la ponderación de los parches) y de huellas dactilares utilizando la regla de suma de resultados.

5.1. Modo identificación

En la Figura 5.1 se muestran los resultados obtenidos, sobre la base TEST, cuando se utiliza el sistema multi-modal para hacer la identificación.

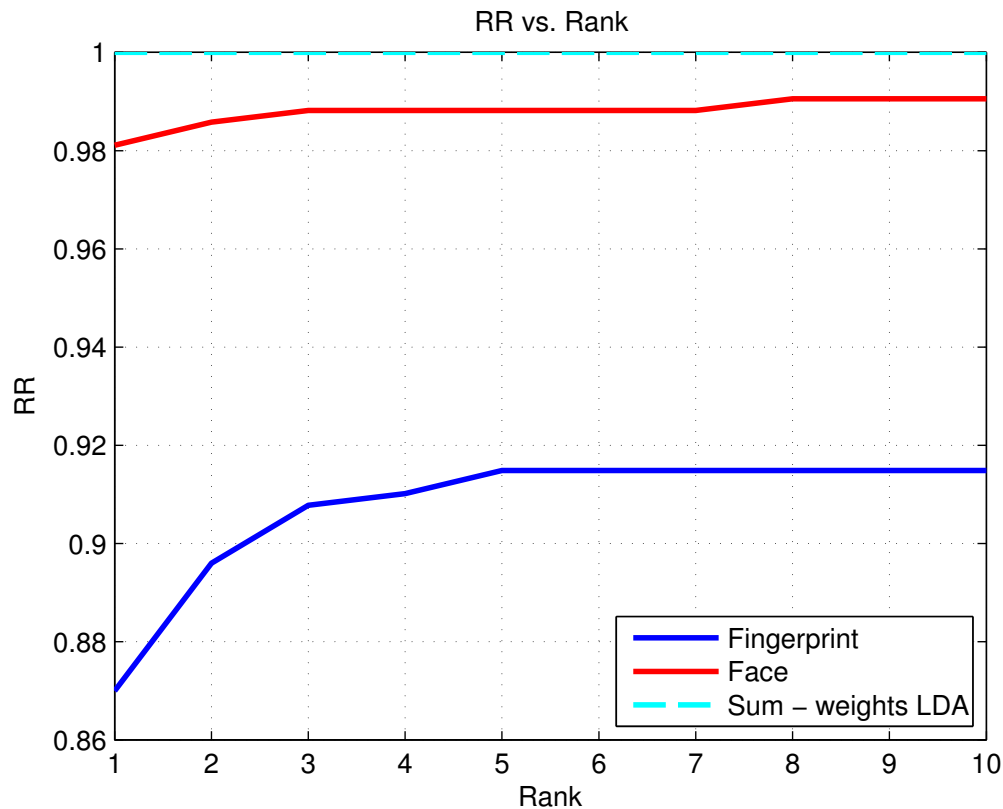


Figura 5.1: RR vs. rank de sistema multi-modal utilizando pesos en parches.

Nuevamente se obtiene al igual que en la base TRAIN, el mejor resultado posible.

5.2. Modo verificación

A continuación se muestran los resultados obtenidos cuando se realiza verificación de identidades. En la Figura 5.2 se muestra el resultado obtenido cuando se utilizan los pesos en parches obtenidos con *LDA* y *SVM* así como el desempeño que se tenía originalmente del sistema de reconocimiento facial. En la Figura 5.3 se muestran los resultados obtenidos por el sistema multi-modal que utiliza las caras y huellas dactilares para realizar la verificación.

En modo verificación el uso de *LDA* vuelve a mejorar el resultado inicial, pero la mejora no es muy grande. Por otro lado, cuando se realiza la fusión con los resultados de las huellas, la mejora es sustancial. En particular para un *FAR* de 10^{-4} la mejora es de un poco más de 10 puntos porcentuales.

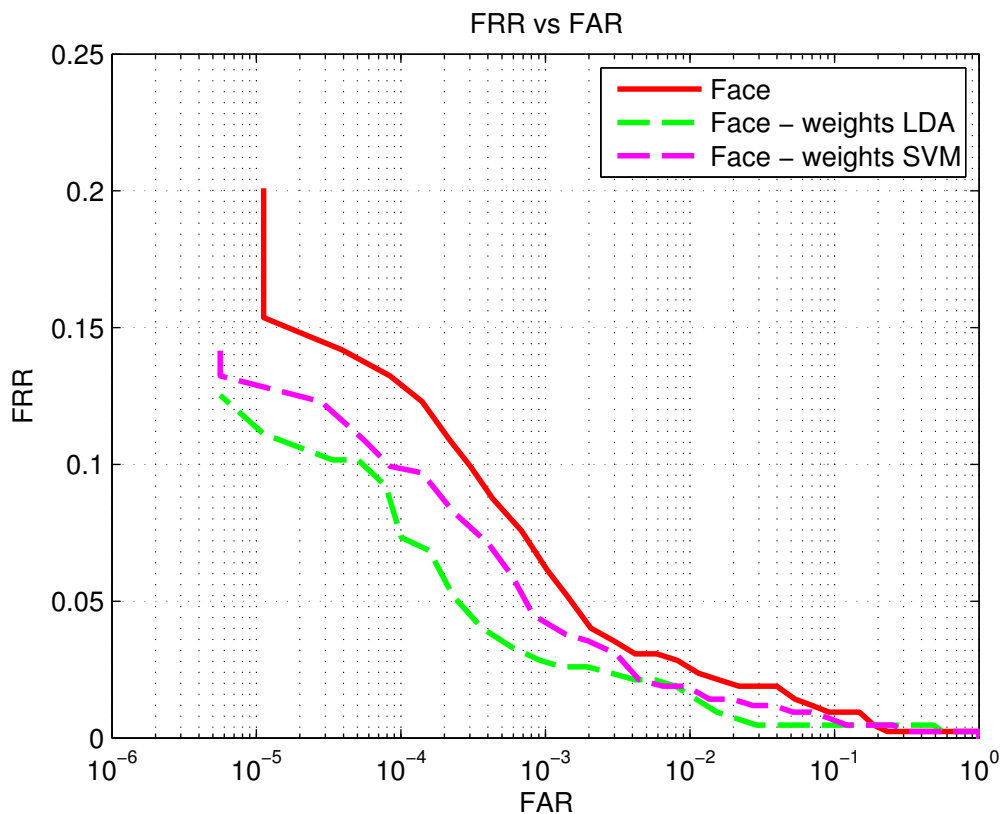


Figura 5.2: FRR vs. FAR para la distintas estrategias de pesos en parches.

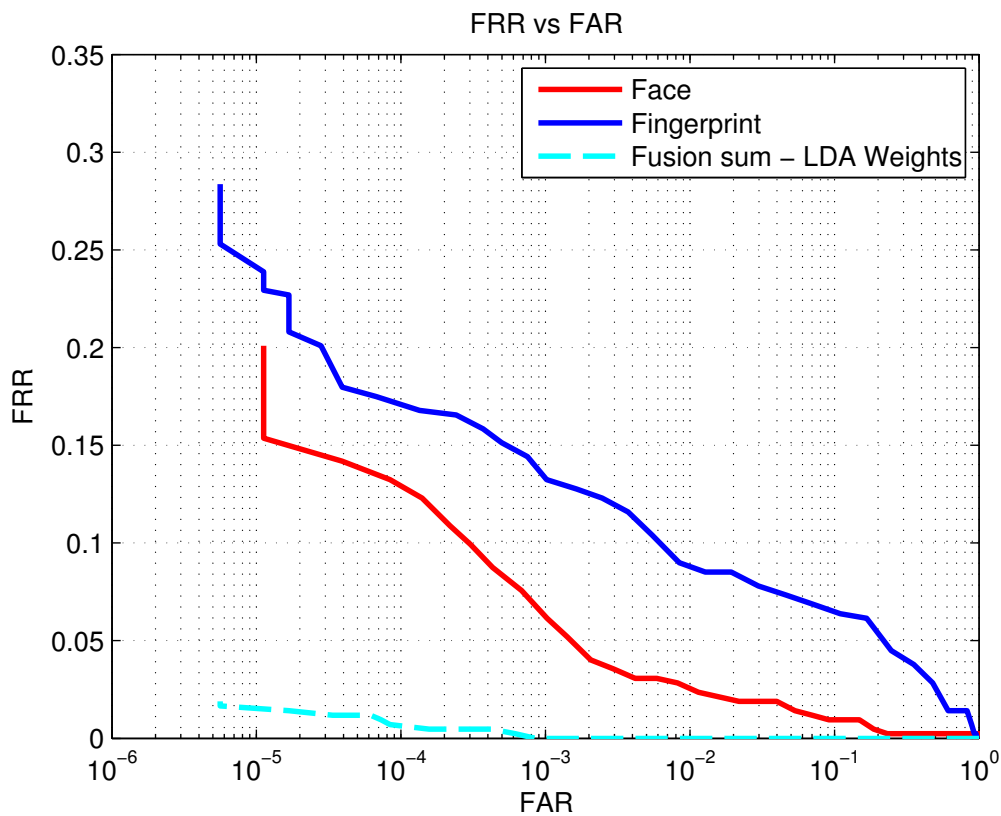


Figura 5.3: FRR vs. FAR de sistema multi-modal utilizando pesos en parches.

Capítulo 6

Conclusiones

En este trabajo se exploró la fusión de clasificadores que provienen de distintas fuentes biométricas utilizando técnicas que se encuentran en la literatura. Para lograr esto se siguieron una serie de pasos que son comunes a toda implementación de un sistema de reconocimiento de patrones: selección de características, normalización de datos y búsqueda de la mejor fusión. Estas técnicas se probaron buscando el mejor desempeño que cada sistema obtenía por separado y, posteriormente, buscando el mejor desempeño de un sistema multi-modal que utiliza al mismo tiempo la información provista por la cara y la huella dactilar de los individuos para realizar la identificación y verificación con una alta precisión. El enfoque utilizado y las técnicas propuestas fueron validadas utilizando para las pruebas de la fusión un subconjunto de las bases de datos originales que se consideraron desconocidas en la etapa de entrenamiento. Al utilizar las estrategias propuestas sobre la base de pruebas se obtuvo una mejora considerable de los resultados tanto en identificación como en verificación. Se destaca que la mejora de los resultados viene dado mayoritariamente por la fusión multi-modal, y no tanto por la mejora de cada sistema inicial.

Como trabajo a futuro se podría pensar en un esquema de selección de características donde esté vinculado el clasificador que se utilizará finalmente. A su vez se puede trabajar con las combinaciones de selección de características junto a fusión de distancias que no fueron contempladas en este trabajo.

Por otro lado se podría trabajar con un conjunto TRAIN y TEST como se hizo en este trabajo, pero utilizando como TRAIN una base de caras completamente distinta a la base de TEST. Esto es de particular interés ya que se lograría encontrar un vector de proyección que sea independiente de la base utilizada. A su vez se podría utilizar bases de datos que no presenten tan buenos resultados inicialmente. En este trabajo no se logró ver la mejoría en identificación cuando fusionan los datos “mejorados” de caras, con los datos de las huellas, ya que la fusión de los datos iniciales daba el mejor resultado posible.

Bibliografía

- [1] T. AHONEN, A. HADID, AND M. PIETIKAINEN. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **28**(12):2037–2041, dec. 2006.
- [2] SEBASTIÁN BERCHESI, LUIS DI MARTINO, AND GABRIEL LEMA. Proyecto faceval, 2012.
- [3] DONG CHEN, XUDONG CAO, FANG WEN, AND JIAN SUN. Blessing of Dimensionality: High-Dimensional Feature and Its Efficient Compression for Face Verification. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3025–3032, June 2013.
- [4] MARIA DE MARSICO, MICHELE NAPPI, DANIEL RICCIO, AND GENOVEFFA TORTORA. NABS: Novel Approaches for Biometric Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **41**(4):481–493, July 2011.
- [5] JOSEF KITTLER, MOHAMAD HATEF, ROBERT P. W. DUIN, AND JIRI MATAS. On combining classifiers. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, **20**:226–239, 1998.
- [6] ARPITA PATRA AND SUKHENDU DAS. Enhancing decision combination of face and fingerprint by exploitation of individual classifier space: An approach to multi-modal biometry. *Pattern Recognition*, **41**(7):2298–2308, July 2008.
- [7] P J PHILLIPS, HYEONJOON MOON HYEONJOON MOON, S A RIZVI, AND P J RAUSS. The FERET evaluation methodology for face-recognition algorithms, 2000.
- [8] K SIMONYAN, OM PARKHI, A VEDALDI, AND A ZISSERMAN. Fisher Vector Faces in the Wild. *Proc. BMVC*, 2013.
- [9] YU SU, SHIGUANG SHAN, XILIN CHEN, AND WEN GAO. Hierarchical ensemble of global and local classifiers for face recognition. *Image Processing, IEEE Transactions on*, **18**(8):1885–1896, aug. 2009.
- [10] YI SUN, XIAOGANG WANG, AND XIAOOU TANG. Deep Learning Face Representation by Joint Identification-Verification. June 2014.