

Sistema Online Basado en Verificación Facial desde Dispositivos Móviles empleando Redes Neuronales Convolucionales

Bruno Tafur Coronel-Zegarra, Marks Arturo Calderón Niquín

11100133@ue.edu.pe, mcladeron@esan.edu.pe

Universidad ESAN, Perú

Resumen: En la presente investigación, se resumen los resultados obtenidos del desarrollo de un sistema online de verificación facial para la seguridad biométrica mediante la utilización de redes neuronales convolucionales. El sistema realiza un proceso de verificación facial, el cual se basa en detectar el rostro, alinearlos, extraer las características y realizar la verificación facial. Asimismo, se utilizan las redes neuronales convolucionales como método para la extracción de vectores de características. Los vectores extraídos mediante la red son representativos de cada rostro y permiten realizar una comparación que devuelva una distancia entre ambos. La red busca maximizar la distancia entre los vectores de rostros distintos y minimizar la distancia entre los de la misma persona. La distancia entre ambos vectores es evaluada tomando en cuenta un punto de corte para determinar si son de la misma persona. En caso de estar sobre el punto de corte, se determina que los rostros son distintos y viceversa. Se describirá la arquitectura del sistema así como las bases teóricas empleadas en el mismo. Se realizaron pruebas de efectividad del sistema en dispositivos móviles con cámara mediante un formulario online con una muestra de 484 comparaciones de 22 personas. Los resultados del prototipo muestran un nivel aceptable confiabilidad.

Palabras clave: convolutional neural networks, computer vision, online systems, face recognition.

Abstract: This paper presents a summary of the results obtained from the development of a face verification system for biometric security using convolutional neural networks. The system follows a process of face verification which detects the face, aligns the face, extracts the features and finally, makes the face verification. As well, it uses convolutional neural networks as a method for feature vector extraction. The obtained vectors using the network are representative of each face and allow to make a comparison that calculates the distance between them. The network aims to maximize the distance between vectors of different faces and minimize it between the ones of the same person. The distance between both vectors is evaluated taking into account a threshold to determine if both photos are of the same person. In case the distance is above the threshold, it is determined that both faces are different and vice versa. The system architecture as well as the theoretical framework will be described. Tests for effectiveness were carried out in mobile devices with camera using an online form with a sample of 484 comparisons with 22 different people. The results of the prototype show an acceptable level of confidence.

Keywords: convolutional neural networks, computer vision, online systems, face recognition.

1. Introducción

La visión computacional es el estudio de los procesos de reconocer y localizar objetos usando el procesamiento de imágenes de tal forma que se logre un mayor entendimiento de estos [1]. Para esto se busca construir tecnologías con dichas capacidades. Asimismo, estos autores plantean que la visión computacional se centra en la extracción de características de la imagen para que estas sean entendidas por una computadora. Mediante el entendimiento de estas características, las computadoras pueden determinar la localización de ciertos objetos dentro de una imagen, reconocerlos, clasificarlos o descomponerlos.

La presente implementación se apoya en el campo de la visión computacional, en especial en los métodos de verificación facial para la creación de un sistema online con dichas funcionalidades. La verificación facial tradicional se basaba principalmente en métodos como *eigenfaces* [13] o más recientemente *Local Binary Patterns* [3] para la extracción de características que faciliten el reconocimiento. Sin embargo, el presente estudio plantea la utilización de redes neuronales convolucionales al ser las que mejores resultados están dando en la actualidad [11].

El proceso a utilizar empieza con la detección del rostro determinando si existe o no una cara y dónde se encuentra dentro de la fotografía. Luego se realiza la alineación del rostro para que se encuentre lo menos rotado posible. A partir de esto, se extraen las características de los rostros alineados y finalmente se logra la verificación facial realizando la comparación de las características de ambas imágenes.

En la presente investigación, las redes neuronales convolucionales permiten extraer características representativas y diferenciales de dos rostros de tal forma que sean comparables entre sí. Esta tiene como salida un vector de características representativo por cada imagen.

Dado que la literatura acerca de verificación facial en dispositivos móviles sobre el esquema de validación de usuarios es carente, se plantea un sistema bajo el uso de redes neuronales convolucionales que presente dichas características. Se busca una implementación que sirva como medida de reducción de riesgo de una suplantación de identidad al momento de realizarse un registro a un servicio en línea. De esta forma el sistema tiene como funcionalidad principal la validación de la foto del Documento Nacional de Identidad (DNI) del ciudadano peruano comparándola con una foto de la misma persona.

El presente artículo está dividido de la siguiente manera. En la Sección 2 se destacan los trabajos previos centrándose en verificación facial y redes neuronales convolucionales. En la Sección 3 se describe la metodología con la teoría que implica la creación del sistema de visión computacional de verificación facial. En la Sección 4 se detalla la implementación describiendo el sistema implementado y su arquitectura. En la Sección 5 se explican los experimentos y resultados y en la Sección 6 la discusión de estos. Finalmente, se presentan las conclusiones del estudio en la Sección 7.

2. Teoría del dominio y trabajos previos

Las implementaciones de verificación facial tradicionalmente se basan en *eigenfaces* o *fisher faces* [13] [2] para la extracción de características. Sin embargo, la superioridad de las redes neuronales convolucionales fue comprobada por la investigación del modelo de DeepFace [11].

En los modelos de redes neuronales convolucionales en la verificación facial destaca principalmente el planteado en FaceNet [9]. Este tiene como principal precedente a DeepFace [11] y Google LeNet [10]. Estos estudios se centraron en la utilización de redes neuronales convolucionales entrenadas con millones de datos de fuentes privadas. Asimismo, recientemente se desarrolló un modelo OpenSource con bases de datos públicas y basado en estos estudios precedentes, este tiene el nombre de OpenFace [1], y demuestra que este tipo de modelos son cada vez más accesibles tanto a usuarios por su movilidad y eficiencia como a desarrolladores.

Por otro lado, dentro de la evaluación de sistemas de seguridad biométrica destaca la realizada en Usability Evaluation of multi-modal biometric verification systems [12], el cual hace una evaluación a detalle de la efectividad, eficiencia y satisfacción del usuario en dichos sistemas y la relación entre ellos. Asimismo, existen investigaciones referentes a la implementación de modelos similares en sistemas online como es el caso de Online Face Recognition System based on Local Binary Patterns and Facial Landmark Tracking [6].

Además, en lo referente al procesamiento y reconocimiento en documentos de identidad o pasaportes, por lo general las investigaciones dan un mayor énfasis al pre-procesamiento de dichos documentos que al reconocimiento [7] [14]. Finalmente, otro uso planteado en documentos de identidad ha sido la detección de documentos falsificados, detectando inconsistencias dentro de este mismo [8].

3. Metodología

La verificación facial se basará en seguir el proceso de detectar los rostros, alinearlos, representarlos y clasificarlos [11] [1]. Asimismo, se utilizará como apoyo el desarrollo OpenSource de OpenFace [1] el cual incluye un modelo pre-entrenado de red neuronal convolucional basado en FaceNet [9].

Una red neuronal convolucional se utiliza para la extracción de características. Es de gran importancia extraer características locales en vez de centrarse en los

pixeles de manera específica en una imagen. Muchos objetos pueden aparecer distorsionados o en distintas posiciones haciendo que sea necesario características generales que describan la imagen en su conjunto o por áreas. Por esta razón, se debe dar importancia a las regiones de la imagen para así detectar características en diversos tamaños y posiciones. Este comportamiento puede ser replicado en una red neuronal forzando a las capas ocultas a combinar fuentes de información local de la imagen. De esta forma, distintas características especiales pueden aparecer en distintos lugares de la imagen y ser detectados de igual manera [5].

Las redes neuronales convolucionales son una extensión de las redes neuronales clásicas pero con más dimensiones al recibir valores matriciales de imágenes en más de un canal. Asimismo, estas redes tienen varias características que las diferencian entre las que resaltan el compartir pesos entre neuronas y el uso de pesos matriciales por cada neurona [4].

Las redes neuronales convolucionales están subdivididas en capas. Las capas más comunes son las siguientes [1]:

- Capas de convolución que deslizan un filtro sobre los valores de características de entrada.
- Capas totalmente conectadas que hallan la sumatoria de los valores de entrada considerando ciertos pesos.
- Capas de submuestreo que suelen obtener el máximo o promedio de regiones espaciales del mapa de características.

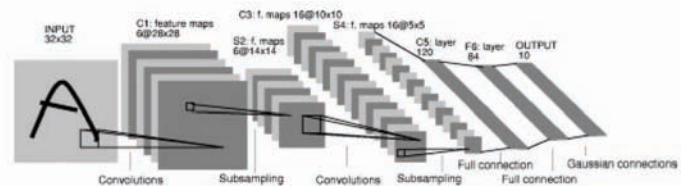


Figura 1: Red neuronal convolucional para reconocimiento de dígitos LeNet-5 [5]

La red planteada devuelve un vector de características por imagen. En la fase de verificación facial, la distancia entre ambos vectores es evaluada tomando en cuenta un punto de corte determinando para establecer si se trata de dos fotos de la misma persona.

4. Implementación

El sistema está dividido en dos módulos principales. El primer módulo representa la página web mediante un prototipo de formulario de registro. El segundo módulo incluye el servicio de verificación facial el cual se encuentra en otro servidor. En conjunto, ambos módulos permiten generar una solución automatizada en la cual solo se requiere, que el usuario suba sus fotos para recibir un resultado.

La decisión de separar el sistema se debe a la conveniencia y orden del desarrollo a futuro. Esto ayuda a evitar recodificaciones y permite la integración sencilla con sistemas existentes. Asimismo, facilita la reutilización. Podría haber más de un sistema web llamando al servicio de verificación facial por separado sin causar interferencias ni reprogramación.

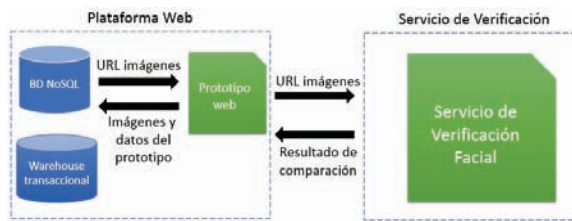


Figura 2: Arquitectura del Sistema

Se ha considerado el Documento Nacional de Identidad (DNI) que se encuentra dentro del contexto peruano. El sistema se podría adecuar a otros contextos en conjunto con el organismo encargado del país donde se desee. El sistema es indiferente a variaciones de tamaño en las imágenes. Sin embargo imágenes más grandes podrían generar una mayor demora en el envío.

4.1. Plataforma Web

El primer bloque de la Figura 2 representa el equivalente a una aplicación web de registro que desee conectarse con el algoritmo. Este incluye la parte visual del prototipo en sí con el formulario para que el usuario interactúe. Además, tiene una base de datos y un data warehouse en la nube.

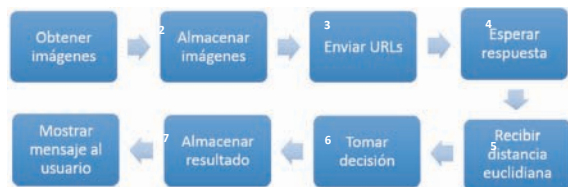


Figura 3: Proceso de Plataforma Web

Esta aplicación recibe los datos del usuario y los almacena en una base de datos no-relacional (NoSQL) en la que no existe una estructura estricta y brinda flexibilidades para almacenamiento de imágenes. Una vez almacenadas las imágenes, el prototipo envía la URL de las imágenes al servicio de verificación facial que se encuentra en otro servidor. Este proceso se ilustra mejor en la Figura 3.

4.2. Servicio de Verificación Facial

El segundo módulo está relacionado con el algoritmo de red neuronal convolucional y se encuentra en otro servidor. Este servidor no está hecho para relacionarse directamente con los usuarios. Este servidor recibe peticiones de la web del primer bloque y responde las consultas automáticamente. A este se le hacen consultas con las URL de dos imágenes faciales para que devuelva resultados de la comparación. Asimismo, este cuenta con la capacidad para recibir peticiones de múltiples páginas web al ser totalmente independiente.

El proceso que se desarrolla en este módulo se explicará a continuación y se resume en la Figura 4.

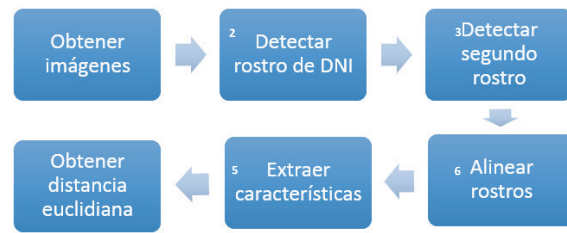


Figura 4: Proceso de Verificación Facial

Obtener imágenes

En primer lugar se obtienen las imágenes. Se recibe la petición con ambas URL. Luego, se extrae cada URL de la petición y se descargan los archivos de cada URL en memoria. Finalmente, se convierte cada archivo en matrices de píxeles para que puedan ser procesados por el resto del servicio.

Detectar Rostro de DNI

La detección del rostro en el DNI es la más compleja dado que generalmente las fotos son tomadas con el DNI sobre una superficie y la foto no brinda información acerca de la orientación de la imagen en estos casos. Es por esto que se debe rotar la imagen hasta encontrar la posición correcta. Utilizando el siguiente procesamiento, las variaciones de color, posición y rotación se pueden dar de manera moderada.



Figura 5: Procesamiento de DNI

Algoritmo 1 Detectar rostro DNI

```

Asegurar: area_dni <= binarizacion(imagen_inicial)
si noHayRostro(area_dni) entonces
    area_dni = transpuesta(area_dni)
si noHayRostro(area_dni) entonces
    rostro <= nulo
sino
    si estaDeCabeza(area_dni) entonces
        voltearDNI(area_dni)
    fin si
    rostro <= recortar(area_dni)
fin si
sino
    si estaDeCabeza(area_dni) entonces
        voltearDNI(area_dni)
    fin si
    rostro <= recortar(area_dni)
fin si

```

El procesamiento del DNI primero detecta el área de la imagen en donde se encuentran píxeles dentro de un rango de colores establecido. Este rango de color

establece la tonalidad de azul que debería tener el DNI. Para esto se utiliza una técnica similar a la binarización con umbral. Se busca los píxeles que se encuentran dentro de un rango de colores en los canales R, G y B. A partir de esto se obtiene una imagen binaria que determina dónde se cumple dicha restricción de color y dónde no. Luego, se le hace un filtrado a la imagen binaria mediante el uso de filtros de 5x5 para suavizar los resultados. Una vez detectada el área, esta imagen binaria, se utiliza para hacer un primer recorte de la imagen.

A partir de esta imagen recortada se busca detectar un rostro dentro de esa zona. La detección del rostro en el DNI se hace usando las características HOG (Histogram of Oriented Gradients) como método de extracción de características y un clasificador SVM (Support Vector Machines). Si no se encuentra el rostro, se procede a obtener la transpuesta de la imagen y se vuelve a buscar el rostro. La transpuesta de una imagen es el equivalente a obtener la transpuesta de la matriz de píxeles de la imagen. Si no se encuentra el rostro, se determina que no existe rostro y se envía mensaje de error al usuario.

Dado que las características HOG en este caso detectan un rostro incluso cuando está rotado 180°, si es que se encuentra el rostro se debe determinar si es que está de cabeza o no. Para esto se compara la distancia entre el rostro detectado y la parte superior del DNI con la distancia con la parte inferior del DNI. En caso que la distancia sea mayor, se voltea la imagen. Una vez que se tiene la imagen en la dirección correcta, se recorta la imagen para extraer solo el rostro agregando un margen de 20 píxeles para que el recorte no sea demasiado cercano a la cabeza de la persona.

Detectar segundo rostro

La detección del rostro de la segunda imagen se hace de manera más sencilla dado que al tomarse la foto de forma vertical, se puede obtener la imagen rotada correctamente sin problemas. Por lo tanto se utilizan de igual manera las características HOG y el clasificador SVM. A partir del rostro detectado, se realiza el recorte. En caso de no encontrar rostro, se devuelve un mensaje de error al usuario.

Alinear imágenes

Una vez extraído el rostro del DNI y el de la segunda foto, se alinea ambas imágenes de tal forma que se pongan derechos los rostros en ojos y nariz y se recorta exactamente la zona del rostro de cada una. En el caso de OpenFace se utiliza solo la alineación afín 2D [1]. Esto se hace buscando que los ojos, nariz y boca aparezcan siempre en posiciones similares.

Extraer características

Posteriormente, las imágenes procesadas son enviadas una por la red neuronal convolucional la cual devuelve un vector de 128 dimensiones para cada imagen. La red neuronal convolucional fue extraída de OpenFace [1], está en su mayoría basada en la red establecida por FaceNet [9]. La estructura se describe en la Tabla 1.

Tabla 1: Estructura de red neuronal convolucional

Capa	Dimensión de entrada	Dimensión de salida
Capa de Convolución (Kernel 7x7x3, 2)	96x96x3	48x48x64
Capa de Normalización	48x48x64	48x48x64
Capa de Submuestreo con Valor Máximo (Kernel 3x3x64,2)	48x48x64	24x24x64
Capa de Normalización	24x24x64	24x24x64
Inception 2	24x24x64	24x24x192
Inception 3 ^a	12x12x192	24x24x256
Inception 3 ^b	12x12x256	12x12x320
Inception 3 ^c	12x12x320	6x6x640
Inception 4 ^a	6x6x640	6x6x640
Inception 4 ^e	6x6x640	3x3x1024
Inception 5 ^a	3x3x1024	3x3x736
Inception 5 ^b	3x3x736	3x3x736
Capa de Submuestreo con valor promedio (Kernel 3x3x1,1)	1x1x736	1x1x736
Capa totalmente conectada	1x1x128	1x1x128
Capa de Normalización	1x1x128	1x1x128

Obtener diferencia de rostros

Una vez obtenidos los resultados de la red neuronal convolucional, se comparan ambos vectores de características obteniendo la distancia euclidiana entre ambos. En este caso, esta se halla mediante la siguiente fórmula:

$$Distancia\ Euclidiana = \sqrt{\sum_{i=1}^n (X_{1i} - X_{2i})^2}$$

Donde X_1 y X_2 son los dos vectores de características, uno por cada rostro y n es el número de elementos de cada vector. En este caso n es 128 dado que la red neuronal convolucional tiene como salida un vector de características de 128 valores por cada rostro. El resultado es enviado de vuelta al prototipo.

El prototipo da una respuesta al usuario dependiendo del resultado y lo almacena en la base de datos para posterior análisis. Para esto, compara la distancia obtenida con el punto de corte. En caso la distancia sea mayor al punto de corte, determina que son rostros de distintas personas y viceversa.

5. Experimentos y resultados

El experimento realizado fue mediante pruebas del sistema por parte de los usuarios. Este fue accedido por los encuestados mediante una URL desde sus dispositivos móviles, desde la cual se les pidió que ingresen sus fotos para la evaluación tal como se muestra en la Figura 6. Para limitar la variabilidad en las muestras, se establecieron instrucciones claras en el formulario para el

envío. Asimismo, se inició con un punto de corte de 0.99 al utilizar el modelo desarrollado de OpenFace [1].

La población del estudio es la población limeña mayor a 18 años con acceso a un *smartphone* e Internet. La muestra de la investigación es de 484 comparaciones de 22 personas.

A continuación se discutirán los resultados encontrados en la investigación. A partir de las imágenes de los rostros, se hizo también una comparación entre el rostro de cada persona con todos los DNIs. De esta forma se obtuvieron resultados para casos en los que se espera una denegación del registro por parte del sistema.

Encuesta - Prototipo de Verificación Facial

Preferible que hagas esta encuesta desde un celular. Es una encuesta para una investigación de la Universidad. No se usarán tus datos para fines fuera de la investigación.

Datos Básicos

Edad:

Género:
 Masculino Femenino

Prototipo

Se probará el funcionamiento de la verificación facial. Ambas fotos las puedes tomar desde tu celular.

Toma una foto de tu DNI:

Foto Ejemplo:

- El DNI debe aparecer completo y estar recto dentro de la foto
- El fondo debe ser de un color diferente y plano, de preferencia un color distinto al azul.

Toma una foto de tu cara:

- Tu cabeza debe aparecer completa en la foto
- Mira directamente a la cámara
- No hacer gestos o ponerse lentes, gorras u otros objetos que pudieran interferir
- El fondo debe ser lo más uniforme posible

SIGUIENTE

Figura 6: Formulario de Plataforma Web

En primer lugar se mostrarán los resultados generales obtenidos en los cuales se consideró el punto de corte en 0.99. El porcentaje total de resultados correctos o exactitud de todos los casos fue de 86% siendo 417 resultados acertados de un total de 484.

Las distancias obtenidas entre todas las comparaciones siguen una distribución similar a la distribución normal. La media se encuentra en 1.32 y la mediana también por lo que los resultados son en su mayoría simétricos.

Para ejemplificar los valores entre grupos, se muestra la distribución de valores agrupados por el valor esperado. Se puede visualizar la diferencia en distribuciones de ambos resultados esperados. Por un lado, las parejas de rostros que se espera que sean rechazados, tienen una distribución que tiene una media en 1.385 y con valores más altos por lo general por encima de 0.99. Por otro lado, las fotos que deberían ser aceptadas tienen una media en 0.777 y con valores por lo general debajo de 0.99.

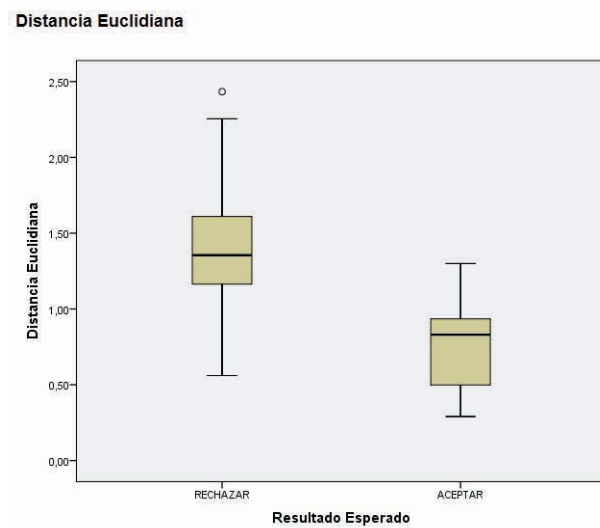


Figura 7: Distribución de distancia euclidiana agrupada por resultado esperado

Asimismo, se realizó un análisis bivariado entre las distribuciones de distancia euclidiana de ambos grupos utilizando el análisis de comparación de medias. Con nivel de confianza de 0.95 (alfa de 0.05) se obtuvo que la significancia es de 0.00, lo cual es menor a alfa (0.05). Por lo tanto, se determinó que los promedios de ambos grupos no son iguales. A partir de esta prueba, se demostró estadísticamente que la distancia euclidiana está relacionada con el resultado esperado.

Tabla 2: Matriz de Consistencia

		Predicción	
		Positivo	Negativo
Real	Positivo	18	4
	Negativo	63	399

De acuerdo a la matriz de consistencia de la Tabla 2, se obtuvo el ratio de falsos positivos y negativos. Según el ratio de falsos positivos, de cada persona que debió ser rechazada por el sistema, el 13.64% fue aceptado como si

fuese la misma persona. Por otro lado, usando el ratio de falsos negativos, de cada persona que debió ser aceptada por el sistema, el 18.18% fue rechazado.

Asimismo, se realizó la curva ROC (Receiver operating characteristic).

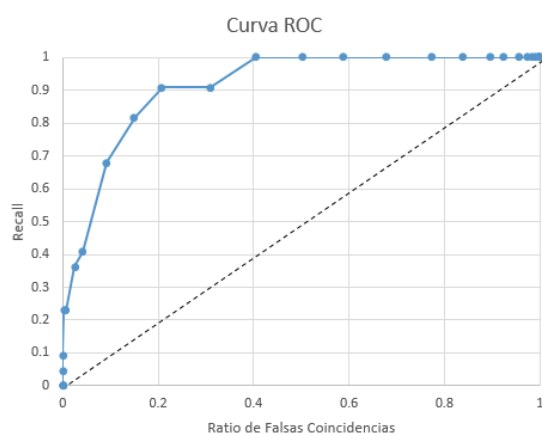


Figura 8: Curva ROC

En la Figura 8 se puede observar el cambio del recall y el Ratio de Falsos Positivos ante distintos puntos de corte. Por lo tanto, el punto de corte óptimo debería estar en la ubicación en la que la curva se acerca lo más posible a la esquina superior izquierda. En este caso ese punto se encuentra cuando el recall está entre 0.8 y 0.9 y el ratio de falsas coincidencias está entre 0.10 y 0.2 aproximadamente. Actualmente, se tiene un recall de 82% y un ratio de falsos positivos de 14%. Por lo tanto, el punto de corte está situado en el rango adecuado.

6. Discusión de los experimentos

A partir de los resultados obtenidos, es posible establecer un punto de corte óptimo al minimizar falsos positivos y falsos negativos. La distancia euclidiana influye drásticamente al momento de tomar la decisión. Esta relación se debe principalmente a un punto de corte establecido. Finalmente, cabe destacar que se obtuvieron resultados aceptables dentro del sistema considerando que solo se contaba con dos muestras de foto por persona de las cuales una era una fotografía extraída a baja resolución del Documento Nacional de Identidad.

7. Conclusiones y trabajos futuros

El presente trabajo ha demostrado la efectividad de un servicio de verificación facial como parte de un sistema móvil online. Este se centró especialmente en el contexto peruano. A partir de esto, se pudo no solo desarrollar un prototipo y una arquitectura, sino también evaluar cómo se relacionaban las partes más importantes del sistema en sus resultados.

En primer lugar, se buscó desarrollar un sistema de visión por computador online mediante redes neuronales convolucionales. Asimismo, el desarrollo buscaba tomar en cuenta los distintos módulos involucrados dentro del sistema y sus conexiones. La presente investigación planteó una arquitectura modular que demostró la separación de funciones como se daría en un caso real.

Además, pudo determinar que se puede obtener una alta efectividad dentro de este tipo de implementaciones.

Asimismo, se plantearon indicadores de efectividad que pueden usarse para evaluar el sistema de forma continua a través del tiempo. Incluso, estos indicadores pueden utilizarse para actualizar la ubicación del punto de corte en tiempo real conforme se van obteniendo mayores resultados. Asimismo, estos indicadores pueden ser evaluados de forma continua dado que la base de datos se actualiza con la información necesaria en el momento del registro de cada nuevo usuario.

En el futuro se podrían considerar y evaluar mejoras en la implementación. Por ejemplo, se podría entrenar una red neuronal convolucional con DNIs con distintas iluminaciones y ángulos para que detecte automáticamente hacia donde rotar el DNI y pueda ser recortado fácilmente. Asimismo, se podría considerar el reentrenamiento de modelos con más fuentes de datos e imágenes de DNIs.

Referencias bibliográficas

- [1]. Amos, B., Ludwiczuk, B., & Satyanarayanan, M. (2016). OpenFace: A general-purpose face recognition library with mobile applications. 1-18.
- [2]. Belhumeur, P.N., Hespanha, J.P., & Kriegman, D.J. 1997. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Patt. Anal. Mach. Intell.* 19, 711-720.
- [3]. Huang, G. B., Lee, H. & Learned-Miller, E. (2012). Learning hierarchical representations for face verification with convolutional deep belief networks. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [4]. Karpathy, A., Johnson, J. & Fei-Fei, L (2016). CS231n Convolutional Neural Networks for Visual Recognition. Stanford University. Extraído desde: <http://cs231n.stanford.edu/>
- [5]. LeCun, Y., Boser, B., Denker, J. S., Howard, R. E., Hubbard, W., & Jacket, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, 1(4), 541-551.
- [6]. Linna, M., Kannala, J., & Rahtu, E. (2015). Online Face Recognition System based on Local. *Advanced Concepts for Intelligent Vision Systems*, 403-414.
- [7]. Ross, A., Jain, A., & Bourlai, T. (2009). On Matching Digital Face Images Against Scanned Passport Photos. *IEEE International Conference on Biometrics, Identity and Security (BldS)*.
- [8]. Sáez-Trigueros, D., Hertlein, H. Meng, L., & Hartnett, M. (2016). Shape and Texture Combined Face Recognition for Detection of Forged ID Documents. *MIPRO 2016*, 1437-1442.
- [9]. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815-823.
- [10]. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Rabinovich, A. (2014). Going Deeper with Convolutions. *ArXiv Preprint arXiv: 1409.4842*, 1-12. <http://doi.org/10.1109/ICCV.2011.6126456>

- [11]. Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1701-1708.
- [12]. Toledano, D. T., Pozo, R. F., Trapote, A. H., & Gómez, L. H. (2006). Usability evaluation of multi-modal biometric. *Interacting with Computers*, 18, 1101-1122.
- [13]. Turk, M. & Pentland, A. (1991). Face Recognition using Eigenfaces. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 586-591.
- [14]. Whitelam, C., Clark, A. D., & Bourlai, T. (2016). Document to Live Facial Identification. In T. Bourlai (Ed.), *Face Recognition Across the Imaging Spectrum* (pp. 223-245). Springer International Publishing.