

Reconocimiento de rostros térmicos usando redes neuronales convolucionales

Raúl Aguilar Figueroa, Giselle Castrejón González, Citlaly Mejía Moreno

Tecnológico Nacional de México/Instituto Tecnológico de Zitácuaro, Departamento de Sistemas y Computación.

Resumen

En este artículo, se presenta un método de reconocimiento de rostros térmicos basado en el uso de redes neuronales convolucionales (CNN's, del inglés Convolutional Neural Networks), un tipo especial de redes neuronales que permite realizar las fases de extracción de características y clasificación. El método propuesto se evalúa bajo las condiciones impuestas por dos experimentos, los cuales se distinguen por emplear los termogramas faciales provenientes de la base de datos Terravic Facial IR Database. Además, para cada experimento, se diseñó un modelo de CNN específico. Particularmente, el Experimento 2 representa uno de los escenarios más complejos para el reconocimiento de rostros térmicos: la obstrucción de parte del rostro debido a la presencia de accesorios faciales. Los resultados experimentales señalan que el uso de CNN's para conducir el proceso de reconocimiento de rostros térmicos, es una opción viable para aumentar la precisión en la categorización de termogramas faciales.

Abstract

This paper presents a thermal face recognition method based on convolutional neural networks (CNN's), a special type of neural network that undertakes the stages of feature extraction and classification. The proposed method is evaluated under the conditions imposed by two experiments, which are distinguished for employ the facial thermograms from the Terravic Facial IR Database. Moreover, for each experiment, a specific CNN model was designed. Particularly, the Experiment 2 represents one of the most complex scenarios for the recognition of thermal faces: the obstruction of part of the face due to the presence of facial accessories. The experimental results point out that the use of CNN's to conduce the thermal face recognition process, is a viable option to increase the accuracy in the categorization of facial thermograms.

Palabras clave: Reconocimiento de Rostros Térmicos, Redes Neuronales Convolucionales, Aprendizaje Profundo.

1. INTRODUCCIÓN

Los sistemas biométricos permiten establecer de forma automatizada la identidad de un individuo a partir de rasgos como las huellas dactilares, el rostro, el iris o la geometría de la mano. En particular, el reconocimiento de rostros se distingue del resto de biometrías por sus cualidades de ser no invasivo y de no requerir de la cooperación del usuario. Actualmente, el despliegue en escenarios de la vida real de aplicaciones basadas en el reconocimiento de rostros, manifiesta una tendencia creciente, y entre algunos de los casos de uso más importantes, se encuentran la búsqueda de niños perdidos [1], el desbloqueo de teléfonos inteligentes [2], el abordaje de vuelos internacionales [3], la detección de sospechosos mediante lentes de sol [4] y el diagnóstico de enfermedades [5].

Tradicionalmente, los métodos de reconocimiento de rostros se han desarrollado basándose en el uso de imágenes capturadas en el espectro visible [6]. No obstante, la precisión de este tipo de métodos se ve afectada por una serie de limitantes, entre las que se encuentran la variación en la iluminación externa, la dificultad en la detección de disfraces faciales, y los cambios en las poses y expresiones faciales [7, 8]. Buscando

superar estos inconvenientes, surge la alternativa de las imágenes faciales obtenidas mediante cámaras infrarrojas [9, 10].

El espectro infrarrojo se divide en cuatro bandas: infrarrojo-cercano ($0.7 - 0.9 \mu m$), infrarrojo de onda corta ($0.9 - 2.4 \mu m$), infrarrojo medio ($3.0 - 5.0 \mu m$), e infrarrojo lejano ($8.0 - 14 \mu m$) [11]. El cuerpo humano emite radiación térmica en las bandas del infrarrojo medio y lejano, las cuales constituyen el espectro infrarrojo térmico [12]. En lo que al rostro se refiere, las imágenes térmicas representan la huella térmica proveniente principalmente de la red vascular subcutánea del rostro, la cual es irreproducible y única para cada persona [13]. Este aspecto es un contraste significativo con respecto a lo que ocurre con las imágenes del espectro visible, las cuales no proporcionan información interna o anatómica del rostro. Además de lo anterior, las imágenes térmicas ofrecen independencia ante la iluminación externa y en general, una mayor robustez ante los factores que usualmente afectan a las imágenes faciales del espectro visible [14, 15]. A pesar de estas ventajas, el reconocimiento de rostros térmicos no está libre de inconvenientes, pues se ve afectado por diversas condiciones internas y externas, incluyendo los cambios en la temperatura ambiente, las variaciones en el metabolismo y el uso de imágenes de prueba obtenidas en un lapso distinto al de las imágenes de entrenamiento [16, 17]. La opacidad del vidrio ante gran parte de la banda infrarroja térmica hace que el uso de anteojos sea otro factor que influya de forma negativa en los termogramas faciales, pues provoca la pérdida de información vascular esencial cercana a los ojos [11]. Además de los anteojos, existen otros accesorios que causan la obstrucción de parte del rostro térmico, como las gorras, pelucas y bigotes falsos.

Buscando mejorar el rendimiento del reconocimiento de rostros térmicos ante los problemas previamente planteados, en el presente artículo se utiliza un modelo inspirado en las redes neuronales biológicas conocido como red neuronal convolucional (CNN, del inglés Convolutional Neural Network), que pertenece al ámbito del aprendizaje profundo. Los modelos basados en el aprendizaje profundo se distinguen por su capacidad para aprender representaciones de los datos crudos de entrada (como los píxeles de una imagen) de una forma automatizada, basándose en una arquitectura compuesta por múltiples capas, en las que, de forma gradual, se generan representaciones cada vez más significativas de los datos de entrada [18, 19]. Este modo de operación de los modelos de aprendizaje profundo, difiere del que presentan los métodos tradicionales de aprendizaje máquina, los cuales están basados en el diseño manual de extractores de características que permitan generar representaciones apropiadas de los datos crudos de entrada, de tal manera que estas representaciones puedan ser procesadas por el método en cuestión [18]. En el caso específico de las CNN's, en los últimos años han alcanzado un éxito sinigual en el área de visión por computadora [20], donde sus aplicaciones van desde la clasificación de partículas elementales en el CERN [21] hasta el diagnóstico de Alzheimer [22].

El estado del arte reporta una implementación escasa de CNN's en el reconocimiento de rostros térmicos. En el trabajo de Simón et al. [23], se realizó la fusión de extractores de características clásicos con una CNN, esto con la finalidad de clasificar imágenes faciales de una base de datos compuesta por imágenes de tipo RGB, depth (mapas de profundidad) y térmicas, donde por cada modalidad de imagen, se tienen tres conjuntos de imágenes que contienen las siguientes variaciones: poses, expresiones e iluminación [24]. Los métodos Local Binary Patterns (LBP), Histograms of Oriented Gradients (HOG), Haar-like rectangular features (HAAR) y Histograms of Gabor Ordinal Measures (HOGOMs), fueron los extractores de características empleados en el estudio. Los resultados experimentales señalaron la conveniencia de fusionar las características generadas por el método HOGOM con aquellas derivadas de la CNN. Usando la misma base de datos que Simón et al., aunque sólo la modalidad térmica, Wu et al. [25] presentaron una CNN capaz de superar el rendimiento de tres métodos de ingeniería de características: LBP, HOG y momentos invariantes, los cuales utilizaron como clasificador el algoritmo de k-nearest neighbor (KNN). En el artículo de Orji et al. [26], se empleó el método de Fast Wavelet-Transform (FWT) para obtener características de los termogramas de la base de datos Terravic Facial IR

Database', después, las características fueron clasificadas con una CNN propuesta por los autores. Con la finalidad de reducir el error cuadrático medio producido por la CNN, se procedió a procesar las características FWT mediante los métodos Principal Component Analysis (PCA) y Linear Discriminant Analysis (LDA), para posteriormente realizar la clasificación con KNN, y finalmente, se combinaron los resultados de clasificación de la CNN con los de KNN. En el trabajo más reciente sobre reconocimiento de rostros térmicos basado en CNN, Sayed et al. [27] propusieron una arquitectura de CNN para clasificar termogramas faciales de la base de datos Arab Open University, los cuales poseen variaciones en posiciones, emociones y condiciones de iluminación. En este trabajo de investigación, se proponen dos arquitecturas de CNN's (CNN-1 y CNN-2) para llevar a cabo el reconocimiento de los termogramas faciales de la base de datos Terravic Facial IR Database, que se caracteriza por contener variaciones en las poses (frontal, izquierda, derecha) y en el ambiente (interior, exterior), además de que los rostros poseen accesorios como gafas y gorras. Para cada modelo de CNN propuesto, se diseñó una configuración experimental específica, de tal forma que el rendimiento de la CNN-1 se analizó con respecto al Experimento 1, mientras que la CNN-2 se evaluó en función del Experimento 2. El Experimento 1 representa el escenario clásico en el cual se evalúan los métodos de reconocimiento de rostros, es decir, los termogramas faciales que componen los conjuntos de entrenamiento y de prueba, son del mismo tipo. No obstante, en el caso del Experimento 2, se estableció un escenario experimental más complejo y demandante: el conjunto de entrenamiento está constituido por muestras de individuos libres de accesorios (lentes y gorras), pero con variaciones en las poses (frontal, izquierda, derecha), mientras que el conjunto de prueba está compuesto exclusivamente de muestras de individuos que portan gafas y que, además, tienen variaciones en las poses. De acuerdo a nuestra revisión del estado del arte, el presente artículo representa el primer esfuerzo en el que un método de reconocimiento de rostros térmicos basado en CNN's, es desplegado en un escenario experimental como el que se plantea en el Experimento 2.

El resto del artículo se organiza de la siguiente manera. En la Sección 2, se presentan algunos conceptos básicos de las CNN's. Las dos arquitecturas de CNN propuestas, se detallan en la Sección 3. El método de reconocimiento de rostros térmicos empleado se expone en la Sección 4. La descripción de los conjuntos de entrenamiento y de prueba de los Experimentos 1-2, al igual que los resultados experimentales, se presentan en la Sección 5. Finalmente, las conclusiones de nuestra investigación se encuentran en la Sección 6.

2. REDES NEURONALES CONVOLUCIONALES: CONCEPTOS BÁSICOS

Una CNN es un tipo de red neuronal especializada en el aprendizaje de características locales de los datos de entrada, los cuales están organizados en múltiples arreglos [18, 19]. Siguiendo los principios de conexiones locales, pesos compartidos, pooling y el uso de una gran cantidad de capas, las CNN's ofrecen robustez ante las variaciones y distorsiones de los datos de entrada, además de una metodología jerárquica espacial para el aprendizaje de patrones [28, 18, 19]. Esta forma de aprendizaje hace posible que las capas de las CNN's generen representaciones cada vez más sofisticadas de los datos de entrada.

En la Figura 1, se muestra una arquitectura típica de CNN. Tradicionalmente, las primeras etapas de un modelo de CNN se componen de una secuencia de capas convolucionales y de capas max-pooling. La salida de la última capa max-pooling se transforma en un vector, el cual se convierte en la entrada de una red neuronal densamente conectada, cuya capa de salida funge como clasificador.

Las capas convolucionales están constituidas por una serie de mapas de características, y a su vez, cada uno de estos mapas está compuesto por unidades (también llamadas neuronas) que tienen asociado un filtro de pesos, de tal manera que todas las unidades que conforman un mapa de características particular emplean un mismo filtro de pesos, mientras que los distintos mapas de características de una capa convolucional utilizan

filtros de pesos diferentes. El filtro de pesos asociado a una unidad de un mapa de características se conecta con una pequeña región de los mapas de características de la capa previa, donde esta región abarca la misma posición de cada mapa de características que conforma la capa previa. La región en cuestión recibe el nombre de mapa receptivo local, y entre dicho mapa receptivo local y el filtro de pesos conectado a él, se produce la operación de convolución, a cuyo resultado se le suma el valor de un bias asociado al filtro de pesos. Posteriormente, esta suma se convierte en la entrada de una función de activación no-lineal. Luego de cada capa convolucional, suele agregarse una capa max-pooling, que tiene la función principal de reducir la dimensión de cada mapa de características de la capa convolucional previa y así, estos mapas de características modificados constituyen la capa max-pooling en cuestión. Las neuronas que definen cada mapa de características de la capa max-pooling, se forman al extraer el mayor valor de una región específica del mapa de características de la capa convolucional previa.

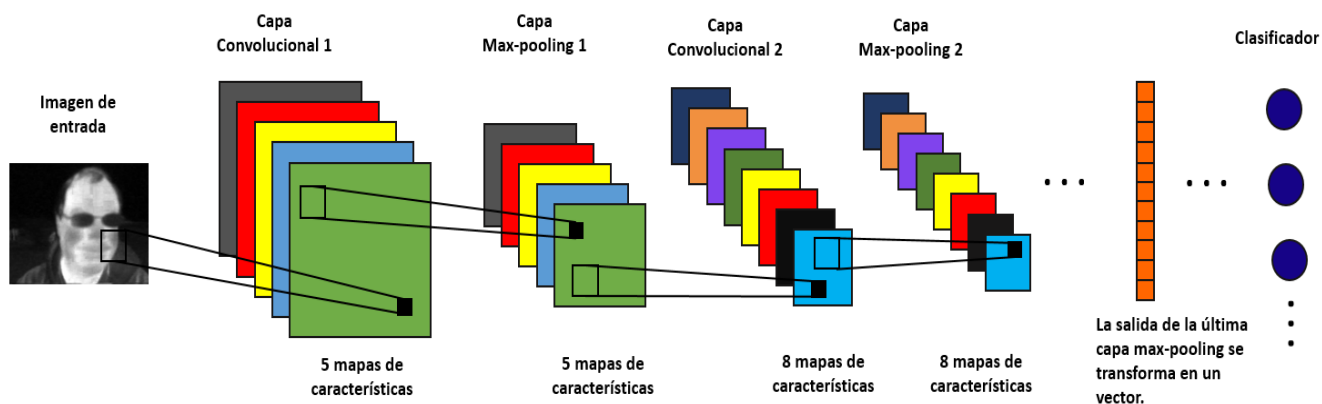


Figura 1. Ejemplo de una arquitectura tradicional de CNN. La imagen de entrada es uno de los termogramas que contiene la base de datos Terravic Facial IR Database.

3. ARQUITECTURAS DE CNN PROPUESTAS

Buscando contribuir en la construcción de sistemas de reconocimiento de rostros térmicos de alto rendimiento, en el presente trabajo de investigación se proponen dos arquitecturas de CNN. En las Subsecciones 3.1-3.2, se ofrece una descripción detallada de dichas arquitecturas.

3.1 Primera Arquitectura de CNN Propuesta (CNN-1)

La primera arquitectura de CNN propuesta, que será denotada como CNN-1, se visualiza en la Figura 2. En total, este modelo se compone de once capas, las cuales incluyen una capa de entrada, tres capas de tipo convolucional (Conv1, Conv2 y Conv3), otras tres capas de tipo max-pooling (MaxPool1, MaxPool2 y MaxPool3), tres capas más de tipo batch normalization (BatchNorm1, BatchNorm2 y BatchNorm3) y finalmente, una capa de salida que categoriza cada uno de los termogramas faciales de entrada (Clasificador Softmax).

La capa de entrada recibe cada uno de los termogramas faciales. Mediante filtros de pesos con dimensiones de 3×3 , las tres capas convolucionales se encargan de realizar el aprendizaje de los patrones locales presentes en las entradas. La capa convolucional Conv1 produce 63 mapas de características, mientras que las capas convolucionales Conv2 y Conv3, generan 100 y 150 mapas de características, respectivamente. En las tres capas convolucionales, se utiliza la función de activación *rectified linear unit* (ReLU) para obtener las neuronas que dan forma a los diferentes mapas de características. La dimensión de los mapas de características de cada una

de las capas convolucionales se reduce mediante las capas max-pooling, que en todos los casos utilizan filtros de 2×2 . Con el objetivo de reducir el sobreajuste de la CNN-1, después de cada capa max-pooling, se añade una capa de tipo batch normalization. En lo que respecta a la capa de salida, está constituida por 5 neuronas, las cuales hacen uso de la función softmax para clasificar los patrones térmicos que proceden de la última capa de tipo batch normalization, BatchNorm3. Los filtros de pesos que se emplean para generar las tres capas convolucionales utilizan un stride de 1 y un padding de 0. Por su parte, los filtros que se usan para construir las tres capas max-pooling, usan un stride de 2 y un padding de 0. Las dimensiones de las salidas de cada capa de la CNN-1, se muestran en la Tabla 1.

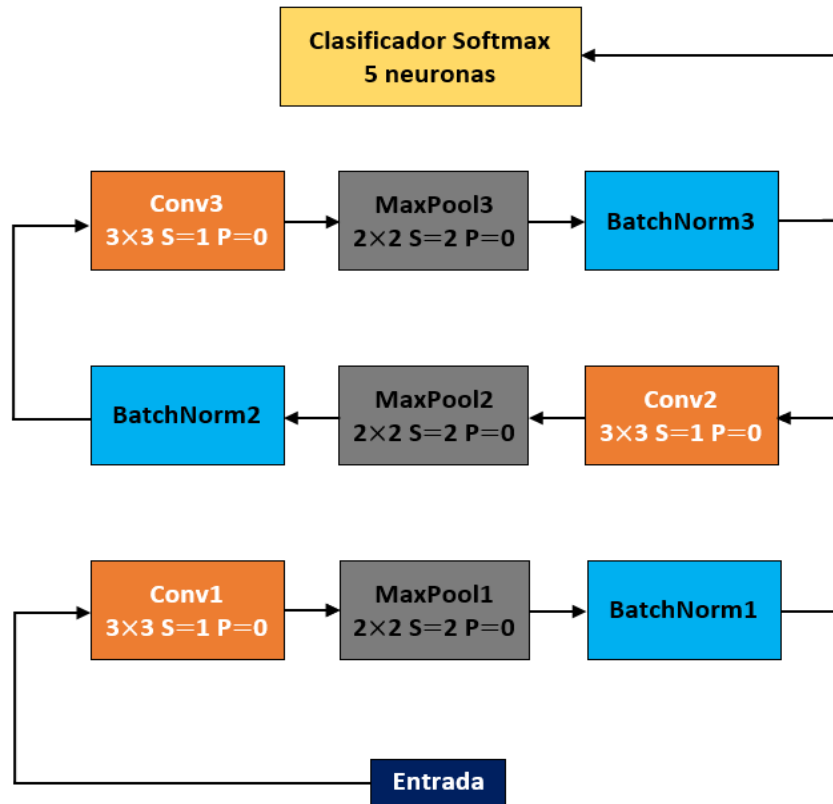


Figura 2. Arquitectura de la primera red neuronal convolucional propuesta (CNN-1)

Capas	Tamaño de salida
Entrada	72×96
Conv1	$63 \times 70 \times 94$
MaxPool1	$63 \times 35 \times 47$
BatchNorm1	$63 \times 35 \times 47$
Conv2	$100 \times 33 \times 45$
MaxPool2	$100 \times 16 \times 22$
BatchNorm2	$100 \times 16 \times 22$
Conv2	$150 \times 14 \times 20$
MaxPool2	$150 \times 7 \times 10$
BatchNorm2	$150 \times 7 \times 10$
Clasificador Softmax	$5 \times 1 \times 1$

Tabla 1. Dimensiones de las salidas de cada una de las capas que conforma la CNN-1

3.2 Segunda Arquitectura de CNN Propuesta (CNN-2)

La segunda arquitectura de CNN propuesta, que será denominada CNN-2, es igual a la CNN-1 en cuanto al número, tipo y disposición de las capas que contiene. Asimismo, la dimensión de los filtros max-pooling, la función de activación, y los valores de padding y stride, son los mismos en las dos CNN's propuestas. La diferencia radica en el número de mapas de características de las capas convolucionales y max-pooling. En la CNN-2, la primera capa convolucional, Conv1, genera 55 mapas de características, mientras que las capas convolucionales Conv2 y Conv3, producen 102 y 145 mapas de características, respectivamente. En la Tabla 2, se indican las dimensiones de las salidas de cada capa de la CNN-2.

Capas	Tamaño de salida
Entrada	72 × 96
Conv1	55 × 70 × 94
MaxPool1	55 × 35 × 47
BatchNorm1	55 × 35 × 47
Conv2	102 × 33 × 45
MaxPool2	102 × 16 × 22
BatchNorm2	102 × 16 × 22
Conv2	145 × 14 × 20
MaxPool2	145 × 7 × 10
BatchNorm2	145 × 7 × 10
Clasificador Softmax	5 × 1 × 1

Tabla 2. Dimensiones de las salidas de cada una de las capas que conforma la CNN-2

4. MÉTODO DE RECONOCIMIENTO DE ROSTROS TÉRMICOS PROPUESTO

Las fases del método propuesto se ilustran en la Figura 3, y su descripción se presenta a continuación:

1. Disminución de la dimensión de los termogramas faciales térmicos que entran a la CNN en cuestión. Los termogramas pasan de un tamaño de 320×240 píxeles a uno de 96×72 píxeles
2. Las propiedades inherentes de la CNN, le permiten llevar a cabo las fases de extracción de características y de clasificación.

El método propuesto y específicamente, las CNN's que se introducen en este artículo, se implementaron en Python 3.5.2 usando Keras [29], que es una librería especializada en el aprendizaje profundo.

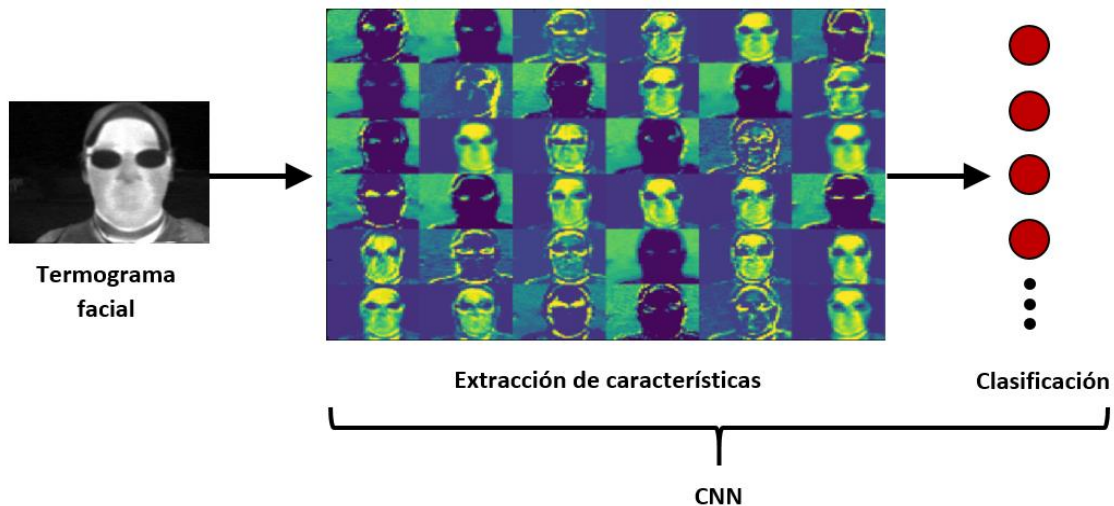


Figura 3. Método de reconocimiento de rostro térmicos propuesto

5. EXPERIMENTOS Y RESULTADOS

5.1 Base de Datos Terravic Facial IR Database

Los termogramas faciales usados en este trabajo, fueron obtenidos de la base de datos Terravic Facial IR Database. Estos termogramas tienen dimensiones de 320×240 , y se caracterizan por poseer variaciones en las poses (frontal, izquierda, derecha) y en el ambiente (interior, exterior), además de que contienen accesorios como gafas y gorras. La base de datos Terravic Facial IR Database es de acceso gratuito y se obtuvo del portal OTCBVS Benchmark Dataset Collection. En dicho portal, se establece que la base de datos en cuestión está compuesta por las muestras del rostro térmico de 20 individuos, no obstante, sólo es posible descargar los termogramas correspondientes a 18 individuos, que en total representan 22,784 imágenes. Es importante especificar que, para fines de este trabajo, se utilizaron únicamente los termogramas pertenecientes a 5 individuos. En la Figura 4, pueden observarse algunas muestras de la base de datos Terravic Facial IR Database.



Figura 4. Muestras de la base de datos Terravic Facial IR Database

5.2 Configuración Experimental

Con el objetivo de implementar el método de reconocimiento de rostros térmicos propuesto, se diseñaron dos escenarios experimentales, los cuales denotaremos como Experimento 1 y Experimento 2. Ambos experimentos se componen de dos conjuntos de termogramas faciales disjuntos: un conjunto de entrenamiento y un conjunto de prueba. El conjunto de entrenamiento se emplea para llevar a cabo la fase de aprendizaje o ajuste de pesos de las dos CNN's propuestas, mientras que el conjunto de prueba se utiliza para evaluar la capacidad de generalización de los modelos de CNN's previamente entrenados. En los dos experimentos propuestos, los termogramas faciales que conforman tanto el conjunto de entrenamiento como el de prueba, fueron seleccionados aleatoriamente y pueden incluir las siguientes variaciones:

- Variación 1: Rostros en diversas poses (frontal, izquierda, derecha).
- Variación 2: Rostros en diversas poses (frontal, izquierda, derecha) con gafas.
- Variación 3: Rostros en diversas poses (frontal, izquierda, derecha) con gorras.
- Variación 4: Rostros en diversas poses (frontal, izquierda, derecha) con lentes y gorras.

En la Tabla 3, se muestran las características de los conjuntos de entrenamiento y de prueba que conforman los Experimentos 1-2. En relación al Experimento 1, las etapas de entrenamiento y de prueba de la CNN-1, se realizan con base a termogramas faciales que se distinguen por poseer las Variaciones 1-4. En el caso del Experimento 2, la CNN-2 es usada para afrontar el problema de la oclusión del rostro térmico debido al uso de accesorios faciales, lo cual provoca una disminución en el rendimiento de los métodos de reconocimiento de rostros térmicos tradicionales. Para abordar apropiadamente el problema planteado por el Experimento 2, el conjunto de entrenamiento se diseñó con termogramas faciales que contienen la Variación 1. Por su parte, el conjunto de prueba se construyó a partir de termogramas faciales que contienen la Variación 2.

ID del Experimento	Conjunto de entrenamiento	Conjunto de prueba
1	Compuesto por 250 imágenes pertenecientes a 5 individuos. Por cada individuo, hay 50 muestras. Contiene las Variaciones 1-4.	Compuesto por 100 imágenes pertenecientes a 5 individuos. Por cada individuo, hay 20 muestras. Contiene las Variaciones 1-4.
2	Compuesto por 250 imágenes pertenecientes a 5 individuos. Por cada individuo, hay 50 muestras. Contiene la Variación 1.	Compuesto por 100 imágenes pertenecientes a 5 individuos. Por cada individuo, hay 20 muestras. Contiene la Variación 2.

Tabla 3. Características de los conjuntos de entrenamiento y de prueba de los Experimentos 1-2.

5.3 Resultados experimentales y discusión

El método de reconocimiento de rostros térmicos propuesto, evaluado en los Experimentos 1-2 a partir de la métrica de precisión rank-1, arrojó los resultados que se exhiben en la Tabla 4.

ID del Experimento	Tasa de reconocimiento rank-1
1	100%
2	92%

Tabla 4. Tasa de reconocimiento rank-1 alcanzado por el método de reconocimiento de rostros térmicos propuesto.

Al analizar los resultados obtenidos al implementar el método propuesto en el Experimento 1, se evidencia que las condiciones de dicha configuración experimental, determinadas por la similitud del tipo de muestras que componen los conjuntos de entrenamiento y validación, facilitan el proceso de reconocimiento por parte de la CNN-1, que alcanza un 100% de precisión en el reconocimiento facial. Estos resultados de clasificación nos indican que, en concordancia con el Experimento 1, la CNN-1 se comporta de forma óptima ante las variaciones en las poses y la presencia de accesorios faciales (Variaciones 1-4).

Por otro lado, el rendimiento del método propuesto en el Experimento 2, en el cual se emplea la CNN-2, es un claro indicativo de la complejidad inherente del problema de oclusión facial causado por el uso de gafas (Variación 2). No obstante, luego de ser entrenada, la CNN-2 tiene una capacidad de generalización aceptable, ya que, a pesar de la pérdida de características vasculares claves en el rostro, alcanza una precisión en la clasificación de 92%. Es importante señalar que las condiciones impuestas al Experimento 2, son únicas en su

clase en la literatura del reconocimiento de rostros térmicos basado en CNN's, por lo cual, el método propuesto representa un paso hacia adelante en la solución al problema de la oclusión facial.

6. CONCLUSIONES Y TRABAJO FUTURO

En este artículo, se introdujo un método de reconocimiento de rostros térmicos basado en el uso de CNN's. Asimismo, se establecieron dos escenarios experimentales (Experimentos 1-2), donde por cada experimento, se propuso una CNN particular. El Experimento 1 representa un esquema tradicional para evaluar el rendimiento de los métodos de reconocimiento de rostros térmicos, mientras que en el Experimento 2 se aborda una de las limitantes esenciales de los termogramas faciales: la oclusión del rostro debido a accesorios como gafas o gorras. En el caso del Experimento 1, la tasa de reconocimiento alcanzada fue de 100%, en tanto que la precisión obtenida en el Experimento 2 fue del 92%. El comportamiento del método propuesto en el Experimento 2, nos permite establecer la hipótesis de que las CNN's pueden ser usadas como alternativas de solución al problema de la oclusión facial, sobre todo si dichas redes son sometidas a un proceso de aprendizaje que involucre un conjunto de entrenamiento con un número elevado de muestras. Con el objetivo de evaluar esta hipótesis en un trabajo futuro, se plantean dos posibilidades: incrementar el número de muestras que conforman el conjunto de entrenamiento del Experimento 2 o escalar este experimento e incluir no sólo un mayor número de muestras por individuo, sino incrementar el número de individuos a ser sometidos al proceso de reconocimiento.

REFERENCIAS

- [1] T. T. India, *Delhi: Facial recognition system helps trace 3,000 missing children in 4 days*, 2018.
- [2] A. Inc., *Usar Face ID en el iPhone o iPad Pro*, 2019.
- [3] J. Airways, *Your Face is Your Boarding Pass: JetBlue Introduces Its First Integrated Biometric Self-Boarding Gate at New York_s John F. Kennedy International Airport*, 2018.
- [4] T. Telegraph, *Chinese police using facial recognition glasses to identify suspects*, 2018.
- [5] Y. Gurovich, Y. Hanani, O. Bar, G. Nadav, N. Fleischer, D. Gelbman, L. Basel-Salmon, P. M. Krawitz, S. B. Kamphausen, M. Zenker, L. M. Bird y K. W. Gripp, «Identifying facial phenotypes of genetic disorders using deep learning,» *Nature Medicine*, vol. 25, pp. 60-64, 2019.
- [6] M. S. Sarfraz y R. Stiefelhagen, «Deep Perceptual Mapping for Cross-Modal Face Recognition,» *International Journal of Computer Vision*, vol. 122, pp. 426-438, 01 5 2017.
- [7] M. K. Bhowmik, K. Saha, S. Majumder, G. Majumder, A. Saha, A. N. Sarma, D. Bhattacharjee, D. K. Basu y M. Nasipuri, «Thermal Infrared Face Recognition _ A Biometric Identification Technique for Robust Security system,» de *Reviews, Refinements and New Ideas in Face Recognition*, P. M. Corcoran, Ed., Rijeka, IntechOpen, 2011.
- [8] R. S. Ghiass, O. Arandjelović, A. Bendada y X. Maldague, «Infrared face recognition: A comprehensive review of methodologies and databases,» *Pattern Recognition*, vol. 47, pp. 2807-2824, 2014.
- [9] R. Aguilar, R. Santiago y J. H. Sossa, «Estudio Comparativo del Reconocimiento de Rostros Térmicos Basado en Características Invariantes,» *Research in Computer Science*, vol. 147, pp. 215-228, 2018.
- [10] V. G. V. Mahesh, A. N. Joseph Raj y P. Arulmozhivarman, «Thermal IR Face Recognition Using Zernike Moments and Multi Layer Perceptron Neural Network (MLPNN) Classifier,» de *Proceedings of the Eighth International Conference on Soft Computing and Pattern Recognition (SoCPaR 2016)*, Cham, 2018.
- [11] S. G. Kong, J. Heo, B. R. Abidi, J. Paik y M. A. Abidi, «Recent Advances in Visual and Infrared Face Recognition: A Review,» *Comput. Vis. Image Underst.*, vol. 97, pp. 103-135, 1 2005.
- [12] I. Pavlidis y P. Symosek, «The imaging issue in an automatic face/disguise detection system,» de *Proceedings IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications (Cat. No.PR00640)*, 2000.
- [13] F. Prokoski, «History, current status, and future of infrared identification,» de *Proceedings IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications (Cat. No.PR00640)*, 2000.

- [14] F. J. Prokoski y R. B. Riedel, «Infrared Identification of Faces and Body Parts,» de *Biometrics: Personal Identification in Networked Society*, A. K. Jain, R. Bolle y S. Pankanti, Edits., Boston, MA: Springer US, 1996, pp. 191-212.
- [15] G. Friedrich y Y. Yeshurun, «Seeing People in the Dark: Face Recognition in Infrared Images,» de *Biologically Motivated Computer Vision*, Berlin, 2002.
- [16] X. Chen, P. J. Flynn y K. W. Bowyer, «IR and Visible Light Face Recognition,» *Comput. Vis. Image Underst.*, vol. 99, pp. 332-358, 9 2005.
- [17] K. R. Kakkirala, S. R. Chalamala y S. K. Jami, «Thermal Infrared Face Recognition: A Review,» de *2017 UKSim-AMSS 19th International Conference on Computer Modelling Simulation (UKSim)*, 2017.
- [18] Y. Lecun, Y. Bengio y G. Hinton, «Deep learning,» *Nature*, vol. 521, pp. 436-444, 2015.
- [19] F. Chollet, *Deep Learning with Python*, 1st ed., Greenwich, CT: Manning Publications Co., 2017.
- [20] A. Krizhevsky, I. Sutskever y G. E. Hinton, «ImageNet Classification with Deep Convolutional Neural Networks,» de *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, USA*, 2012.
- [21] «Quark versus Gluon Jet Tagging Using Jet Images with the ATLAS Detector,» Geneva, 2017.
- [22] M. Liu, D. Cheng, K. Wang, Y. Wang y Alzheimer's Disease Neuroimaging Initiative, «Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis,» *Neuroinformatics*, vol. 16, pp. 295-308, 01 10 2018.
- [23] M. O. Simón, C. Corneanu, K. Nasrollahi, O. Nikisins, S. Escalera, Y. Sun, H. Li, Z. Sun, T. B. Moeslund y M. Greitans, «Improved RGB-D-T based face recognition,» *IET Biometrics*, vol. 5, pp. 297-303, 2016.
- [24] O. Nikisins, K. Nasrollahi, M. Greitans y T. B. Moeslund, «RGB-D-T Based Face Recognition,» de *2014 22nd International Conference on Pattern Recognition*, 2014.
- [25] Z. Wu, M. Peng y T. Chen, «Thermal face recognition using convolutional neural network,» de *2016 International Conference on Optoelectronics and Image Processing (ICOIP)*, 2016.
- [26] C. Orji, E. Hurwitz y A. Hasan, «Thermal Imaging Using CNN and KNN Classifiers with FWT, PCA and LDA Algorithms,» *Seventh International Conference on Computer Science, Engineering and Information Technology (CCSEIT 2017)*, p. 133/143, 2017.
- [27] M. Sayed y F. Baker, «Thermal Face Authentication with Convolutional Neural Network,» *Journal of Computer Science*, vol. 14, 12 2018.
- [28] Y. Lecun, L. Bottou, Y. Bengio y P. Haffner, «Gradient-based learning applied to document recognition,» *Proceedings of the IEEE*, vol. 86, pp. 2278-2324, 11 1998.
- [29] F. Chollet y others, *Keras*, 2015.

Correo electrónico autor: raul.aguilar.figueroa@gmail.com