

Arquitectura mini-Xception para reconocimiento de sexo con rostros mestizos del norte del Perú

Mini-Xception Architecture for sex recognition using non-Caucasian faces from the north of Peru

Dr. Iván Eduardo Salvador Briceño¹
Dr. Guillermo Guerrero Ojeda²
Ing. Gloria Salvador Vásquez³

RESUMEN

La Arquitectura mini-Xception desarrollada para reconocimiento de sexo de personas fue aplicada a rostros caucásicos, en la presente investigación se aplicó en la identificación de sexo a partir de una serie de fotos, tamaño carnet a colores de estudiantes de la Universidad Nacional Pedro Ruiz Gallo de Lambayeque, con el propósito de probar la eficacia del reconocimiento de la citada arquitectura para personas de rasgos no caucásicos. El conjunto de fotos se uniformizó en cuanto a dimensiones, se modificaron las que tenían la cara desproporcionada, descartándose las muy pequeñas y las de personas que tenían cerquillo hasta los ojos o figuras de caras en la vestimenta. Toda esta labor se hizo por inspección visual, con ayuda de la misma arquitectura y algunos programas de edición de fotos, como Photoshop. La identificación correcta alcanzó el 57% en comparación del 96% logrado en rostros caucásicos.

Palabras clave: visión por computadora, reconocimiento facial, Python, red neuronal de convolución.

ABSTRACT

The Mini-Xception Architecture developed for the recognition of the sex of people was applied to Caucasian faces. In this research, it was applied in the identification of sex from a series of color passport-size photos of students from the Pedro Ruiz Gallo National University from Lambayeque, in order to test the effectiveness of the recognition of the aforementioned architecture for people with non-Caucasian features. The set of photos was standardized in terms of dimensions, those with disproportionate faces were modified, discarding the very small ones and those of people who had eye-catches or face figures in the clothing. All this work was done by visual inspection, with the help of the same architecture and some photo editing programs, such as Photoshop. Correct identification reached 57% compared to 96% in Caucasian faces.

Keywords: computer vision, facial recognition, Python, convolution neural network.

1 Universidad Nacional Pedro Ruiz Gallo, Departamento Académico de Ingeniería en Computación e Informática. isalvador@unprg.edu.pe

2 Universidad Nacional Pedro Ruiz Gallo, Departamento Académico de Ingeniería en Computación e Informática. gguerrero@unprg.edu.pe

3 Universidad Nacional Pedro Ruiz Gallo, Escuela Profesional de Ingeniería de Sistemas. gsalvador22@unprg.edu.pe

INTRODUCCIÓN

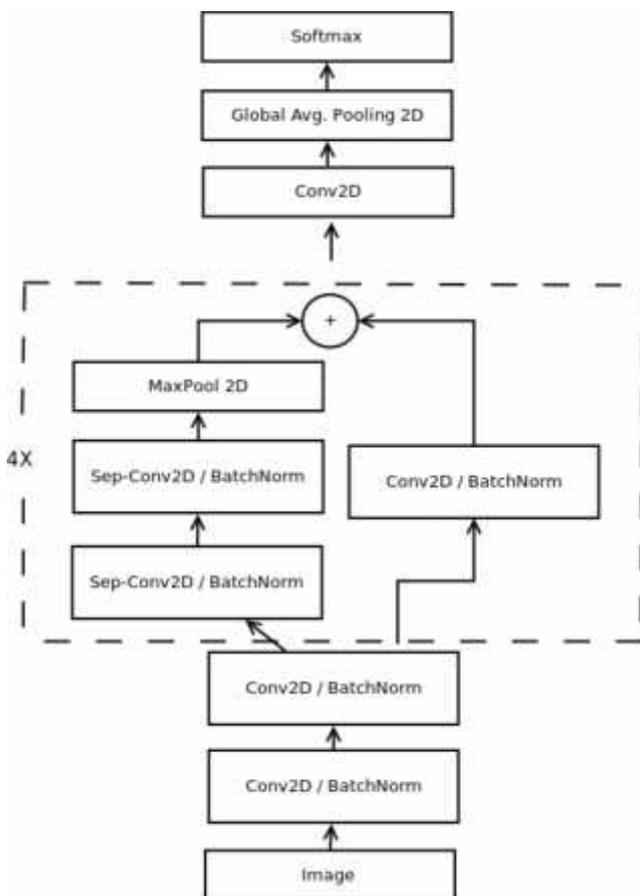
Una tecnología de la inteligencia artificial, es la robótica que pretende dotar de inteligencia a una máquina que sea capaz de ayudar en ciertas tareas a las personas a nivel doméstico y en otros niveles, como el deporte; y para poder ayudar debe ser capaz de interactuar con las personas como si fuera otra persona. La etapa inicial de la interacción es reconocer a la persona a la que servirá, esto se logra primero en determinar el sexo de la persona y de acuerdo a ello actuar.

Existen varios trabajos de investigación para el reconocimiento de imágenes (Simony, 2014), (Szegedy, 2016), en nuestra investigación hemos elegido el que han propuesto Arriaga et al. (2017) que reconoce sexo de imágenes de hombre y mujeres; y como ellos proponen probar su trabajo de redes neuronales (Calvo, 2015), (Lujan, 2016) con rostros no caucásicos para ver como varía la eficiencia del reconocimiento, ya que ellos han probado con rostros caucásicos en su mayoría personajes conocidos en los medios de comunicación. Por eso, nosotros vamos a probar con rostros no caucásicos.

Antecedentes

Según Arriaga et al. (2017) las redes neuronales de convolución (CNN) se usan comúnmente para extracción de características que incluyen un grupo de capas totalmente conectadas para tal fin. Las capas completamente conectadas tienden a mantener muchos de sus parámetros en este tipo de redes neuronales. Específicamente como mencionan el trabajo de Simonyan & Zisserman (2014), estas redes contienen aproximadamente 90% de todos sus parámetros en sus capas finales conectadas. Asimismo, mencionan el trabajo arquitectónico sobre inyección descrito en el trabajo de Szegedy et al. (2016) donde se reduce la cantidad de parámetros en sus últimas capas incluyendo una operación global de Pooling. Esta reducción se logra en cada mapa de características tomando el promedio sobre todos los elementos en el mapa de características y usando la convolución separada en profundidad que reduce mucho más la cantidad de parámetros separando el proceso de la extracción de características y su combinación dentro de su capa de convolución. El modelo de Arriaga et al. (2017) toma esos avances y ésta es su descripción gráfica, figura 1.

Figura 1
 Modelo de arquitectura propuesta para clasificación de rostros (Tomado de Arriaga et al. (2017))



MÉTODOS Y MATERIALES

Se pidió una muestra de fotos de postulantes a la universidad de fotos, se nos alcanzó solo fotos, pero no dato sobre su sexo, así que tuvimos que pedir a otra oficina fotos de estudiantes las que vinieron con los datos de su sexo. Dada la variedad de dimensiones y características de las imágenes, tuvimos que hacer una inspección visual y con ayuda de software de edición de fotos y las muy inapropiadas para usarlas las descartamos, por dimensiones muy reducidas o fotos donde no se identificaba una cara por el software debido a cerquillos pronunciados. El número de fotos usadas, finalmente, fueron 4149 siendo que se nos entregó 4163 fotos.

El lenguaje de programación que se usa en la arquitectura de Arriaga et al. (2017) es Python, y usamos la versión 3.5 del entorno de desarrollo (IDLE) proporcionado por la organización que mantiene el lenguaje; en una computadora con el procesador Intel i5 y sistema operativo Windows 10 Enterprise, con su versión de noviembre 2019, memoria de acceso aleatorio (RAM) de 8 Gigabytes. El programa se modificó para que tomara en forma automática todas las fotos y

analizara cada una de las 4149 fotos, y en la computadora que usamos el análisis y clasificación de todas ellas tomo un tiempo aproximado de 8 minutos. Y con los resultados de clasificación que daba el sexo por la arquitectura se contrasto con los datos entregados donde se decía el sexo de cada estudiante usando un programa de hoja de cálculo, en este caso Excel en versión 2016.

RESULTADOS

El resultado arrojado por la arquitectura se contrasto con los datos entregados y se muestra a continuación en la tabla 1 y en figura 2.

Tabla 1:

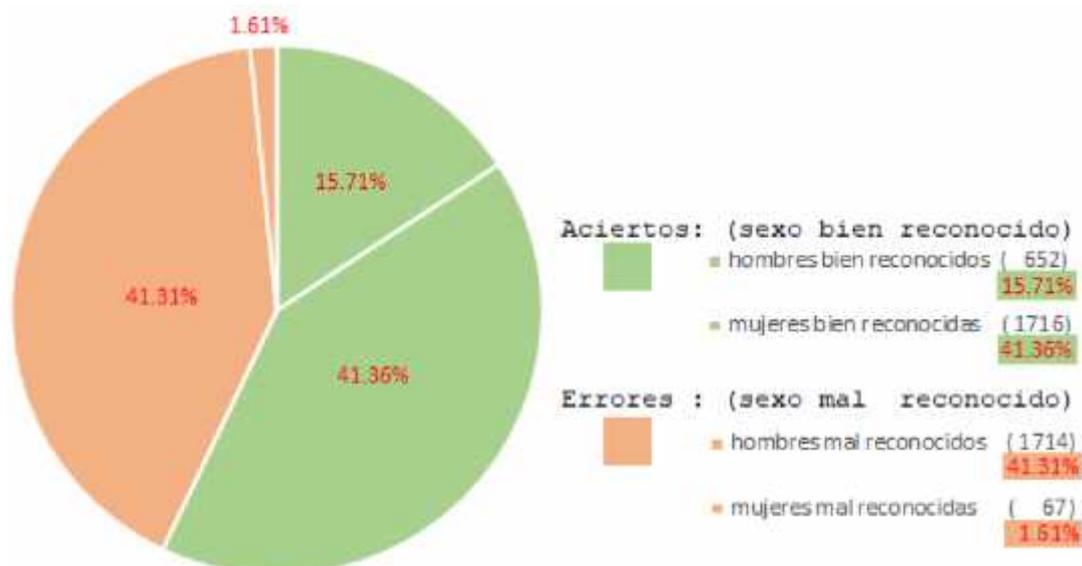
Resultados obtenidos con el uso de la arquitectura de Arriaga et al. (2017) con fotos de personas mestizas

Hombres	2366	% hombres	57.03
Mujeres	1783	% mujeres	42.97
TOTAL	4149		
Aciertos (sexo bien reconocido)	2368	% acierto	57.07
Hombres bien reconocidos	652		27.53
Mujeres bien reconocidas	1716		72.47
Errores (sexo mal reconocido)	1781	% errores	42.93
Hombres mal reconocidos	1714		96.24
Mujeres mal reconocidas	67		3.76

Fuente Elaboración propia de los autores.

Figura 1

Resultados del reconocimiento del sexo por el algoritmo de Arriaga et al. con rostros no caucásicos.



Nota Autoría propia.

DISCUSIÓN

Comparando con el trabajo de investigación de Arriaga et al. (2017) donde se obtuvo el 96% de aciertos del reconocimiento, donde se entrenó a la red neuronal de convolución con personas de rasgos caucásicos y también se le probó con este mismo tipo de personas, pero en nuestro trabajo de investigación se obtuvo como resultado el 57% de aciertos del reconocimiento, habiendo una diferencia notable de casi 40%, donde aplicamos el mismo entrenamiento con rostros caucásicos, pero se probó con rostros mestizos del norte del Perú, específicamente estudiantes de la universidad nacional Pedro Ruiz Gallo que se encuentra en la ciudad de Lambayeque.

Analizando un poco más los resultados, podemos tratar de explicar el mal reconocimiento respecto de los hombres en tal alto porcentaje (41.31%), quizá se deba a la juventud de esas personas, que se encuentran entre los 17 a 24 años de edad.

CONCLUSIONES

Podemos concluir que esta arquitectura, como casi todas, depende del entrenamiento que reciba y por eso cuando se le prueba con rostros mestizos del norte del Perú tiene un bajo porcentaje de aciertos en el reconocimiento del sexo en imágenes fijas.

También es necesario mencionar que en el transcurso de la realización de esta investigación hemos encontrado otras arquitecturas, por ejemplo, FACENET de Brownlee (2019), que deben ser probadas para ver su eficacia y poder elegir la más apropiada para el reconocimiento del sexo analizando caras.

El nivel de eficacia logrado con la arquitectura mini-Xception para reconocimiento de sexo con rostros mestizos del norte del Perú fue del 56%. Los niveles más altos de eficacia de la arquitectura mini-Xception en el reconocimiento sexual en rostros no caucásicos depende del entrenamiento.

REFERENCIAS BIBLIOGRÁFICAS

- Arriaga, Octavio; Paul G. Plöger y Matias Valdenegro (2017). Real-time Convolutional Neural Networks for Emotion and Gender Classification, b-it Bonn-Aachen International Center for Information Technology and the Hochschule Bonn-Rhein-Sieg.
- Brownlee, Jason (2019). Deep Learning for Computer Vision, Image Classification, Object Detection and Face Recognition in Python. Machine Learning Mastery. USA.
- Calvo, Diego (2015). Red neuronal convolucional. Accesado el día 21 febrero de 2019, desde <http://www.diegocalvo.es/red-reuronal-convolucional.html>.
- Lujan, Guillermo. (2016). Las redes neuronales. Accesado el día 13 de diciembre de 2018, desde <https://www.xataka.com/robotica-e-ia/las-redes-neuronales-que-son-y-por-que-estan-volviendo.html>.
- Simonyan, Karen and Andrew Zisserman (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv.1409.1556.
- Szegedy, Cristian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wognna (2016). Rethinking the inception architecture for computer vision, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Pages 2818-2826.