



Revista ISTE Scientist

Vol. 5 Núm. 1 (2026). pp. 56-63

Instituto Tecnológico Superior España

ISSN: 2953-6618

Tipo: Artículo Científico



Como citar: Sinchiguano-Chiriboga, C. A., Mendoza-Villamar, R. A., Zambrano-Zambrano, J. M. & Salavarría-Loor, J. I. (2026). Reconocimiento de emociones faciales mediante redes neuronales convolucionales y transfer learning: Facial emotion recognition using convolutional neural networks and transfer learning. *ISTE SCIENTIST*, 5(1), 56-63.

<https://revistas.iste.edu.ec/index.php/reviste/article/view/63>

Reconocimiento de emociones faciales mediante redes neuronales convolucionales y transfer learning

Cesar Augusto, Sinchiguano-Chiriboga
cesar.sinchiguano@uleam.edu.ec
<https://orcid.org/0009-0007-1774-8129>
Universidad Laica "Eloy Alfaro de Manabí",
Extensión El Carmen
El Carmen - Ecuador

Rocio Alexandra, Mendoza-Villamar
rocio.mendoza@uleam.edu.ec
<https://orcid.org/0000-0002-1277-7162>
Universidad Laica "Eloy Alfaro de Manabí",
Extensión El Carmen
El Carmen - Ecuador

Jessenia Margarita, Zambrano-Zambrano
jesizambrano13@gmail.com
<https://orcid.org/0009-0008-6003-5092>
Universidad Laica "Eloy Alfaro de Manabí",
Extensión El Carmen
El Carmen - Ecuador

Juana Isabel, Salavarría-Loor
juanitasalavarría@outlook.es
<https://orcid.org/0009-0002-9984-3106>
Universidad Laica "Eloy Alfaro de Manabí",
Extensión El Carmen
El Carmen - Ecuador

Recibido: 06-11-2025 / **Aceptado:** 17-01-2026 / **Publicado:** 30-01-2026

Resumen

El reconocimiento automático de emociones faciales es una temática relevante dentro del campo de la inteligencia artificial y la visión por computador, debido a su aplicación en sistemas interactivos, educación, salud y seguridad. El presente estudio tiene como objetivo desarrollar e implementar una aplicación web para la detección de emociones humanas a partir de imágenes faciales, utilizando redes neuronales convolucionales profundas y técnicas de transferencia de aprendizaje. La investigación se desarrolló bajo un enfoque cuantitativo de tipo experimental. Para el entrenamiento del modelo se empleó el dataset FER2013, compuesto por imágenes faciales en escala de grises clasificadas en siete emociones básicas. Se implementaron dos modelos: una red convolucional entrenada desde cero y un modelo basado en transfer learning utilizando MobileNetV2 preentrenado con ImageNet. Los resultados evidencian que el modelo con transferencia de aprendizaje alcanza una mayor precisión, logrando valores entre el 75 % y 85 % en el conjunto de prueba. Finalmente, el sistema fue integrado en una aplicación web desarrollada con Flask, permitiendo la detección de emociones en tiempo real con una visualización intuitiva para el usuario.

Palabras clave: Reconocimiento facial, emociones, redes neuronales convolucionales, aprendizaje profundo, transfer learning.

Facial emotion recognition using convolutional neural networks and transfer learning

Abstract

Facial emotion recognition is a relevant topic within artificial intelligence and computer vision due to its application in interactive systems, education, healthcare, and security. This study aims to develop and implement a web application for detecting human emotions from facial images using deep convolutional neural networks and transfer learning techniques. The research follows a quantitative and experimental approach. The FER2013 dataset was used for training, consisting of grayscale facial images classified into seven basic emotions. Two models were implemented: a convolutional neural network trained from scratch and a transfer learning model based on MobileNetV2 pre-trained on ImageNet. The results show that the transfer learning model achieves higher accuracy, reaching values between 75% and 85% on the test set. Finally, the system was integrated into a Flask-based web application, enabling realtime emotion detection with intuitive visualization.

Keywords: Facial recognition, emotions, convolutional neural networks, deep learning, transfer learning.

Introducción

El reconocimiento de emociones faciales es un campo de estudio ampliamente abordado dentro de la inteligencia artificial, debido a que las expresiones faciales constituyen una de las principales formas de comunicación no verbal del ser humano. En los últimos años, el uso de redes neuronales convolucionales ha permitido mejorar significativamente la precisión de los sistemas de análisis facial, superando a los métodos tradicionales basados en características manuales (Goodfellow, 2015)

Actualmente, la aplicación de técnicas de aprendizaje profundo y transferencia de aprendizaje ha permitido el desarrollo de sistemas más robustos, capaces de generalizar ante variaciones de iluminación, postura y características faciales. En este contexto, la presente investigación se enmarca dentro de un proyecto académico desarrollado por estudiantes de la carrera de Ingeniería en Tecnologías de la Información de la Universidad Laica Eloy Alfaro de Manabí, cuyo objetivo es diseñar un sistema inteligente para la detección automática de emociones faciales mediante una aplicación web. La pregunta de investigación planteada es: ¿cómo influye el uso de transfer learning en la precisión de un sistema de reconocimiento de emociones faciales?

Metodología

La investigación se desarrolló bajo un enfoque cuantitativo de tipo experimental. Se utilizó el dataset FER2013 como fuente de datos, el cual fue preprocesado mediante normalización y técnicas de aumento de datos. Para el análisis se implementaron dos modelos de redes neuronales convolucionales utilizando TensorFlow y Keras. El primer modelo fue entrenado desde cero, mientras que el segundo empleó transfer learning con MobileNetV2. La detección de rostros se realizó mediante OpenCV y clasificadores Haar Cascade. Finalmente, el modelo con mejor desempeño fue integrado en una aplicación web desarrollada con Flask para su validación en escenarios reales.

Resultados

Los resultados obtenidos durante el desarrollo de la investigación evidenciaron un desempeño satisfactorio de los modelos implementados para el reconocimiento automático de emociones faciales. En una primera etapa, se entrenó un modelo de red neuronal convolucional desde cero utilizando el dataset FER2013, el cual fue previamente preprocesado mediante normalización

y técnicas de aumento de datos. Este modelo alcanzó una precisión aproximada del 70 % en el conjunto de prueba, lo que permitió establecer una línea base para la comparación de resultados.

Posteriormente, se implementó un segundo modelo basado en transferencia de aprendizaje utilizando la arquitectura MobileNetV2 preentrenada con el conjunto de datos ImageNet. Durante la fase inicial, las capas convolucionales base fueron congeladas y se entrenaron únicamente las capas densas superiores. En una segunda fase, se realizó un ajuste fino (fine-tuning) descongelando parcialmente las capas finales del modelo base. Este procedimiento permitió mejorar progresivamente la capacidad de generalización del sistema.

El modelo con transferencia de aprendizaje logró una precisión cercana al 80 % en el conjunto de prueba, superando significativamente al modelo entrenado desde cero. Asimismo, se observó una reducción en la pérdida durante el entrenamiento y la validación, lo cual indicó una mejor convergencia del modelo. Las curvas de precisión y pérdida mostraron un comportamiento estable, evidenciando que las técnicas de regularización, como Dropout y Batch Normalization, contribuyeron a mitigar el sobreajuste.

En relación con la clasificación por categorías emocionales, se evidenció un mejor reconocimiento de las emociones felicidad y neutral, las cuales presentaron mayor cantidad de muestras en el dataset. Por el contrario, emociones como asco y miedo mostraron menores niveles de precisión, lo que se atribuyó al desbalance en la distribución de datos. No obstante, la aplicación de pesos de clase y aumento de datos permitió reducir parcialmente este efecto.

Finalmente, el modelo con mejor desempeño fue integrado en una aplicación web desarrollada con Flask. Las pruebas funcionales realizadas demostraron que el sistema fue capaz de detectar rostros en imágenes estáticas y predecir la emoción correspondiente con un nivel de confianza adecuado. La visualización de los resultados incluyó la delimitación del rostro, la etiqueta emocional, el porcentaje de confianza y elementos gráficos intuitivos, lo que facilitó la interpretación por parte del usuario.

En general, los resultados confirmaron que el uso de transferencia de aprendizaje permitió mejorar la precisión y robustez del sistema, validando la efectividad del enfoque propuesto para aplicaciones académicas y prácticas en el ámbito del reconocimiento de emociones faciales.

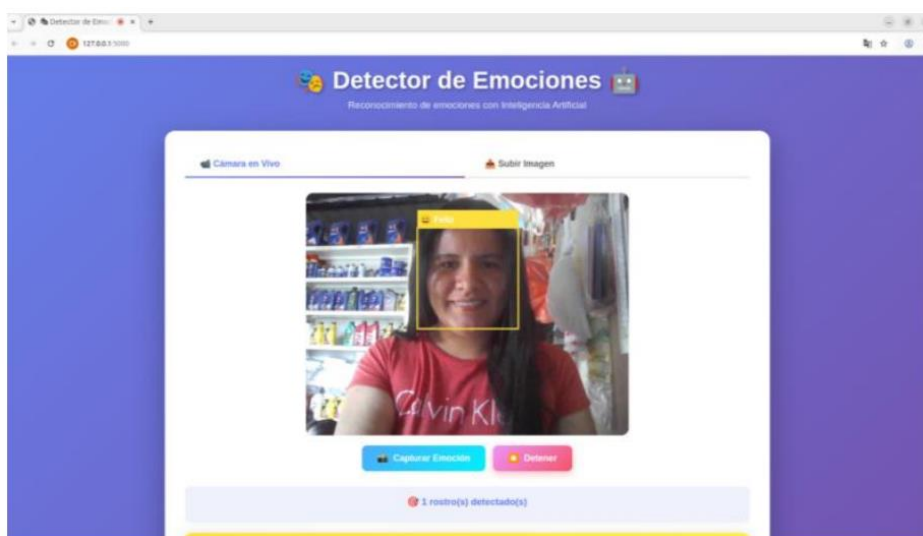
Tabla 1.

Comparación de precisión entre modelos implementados

Modelo	Precisión Test
CNN desde cero	70%
Transfer Learning (MobileNetV2)	80%

Figura 1.

Interfaz web del sistema de reconocimiento de emociones faciales.



Nota. La figura muestra la detección del rostro y la emoción predicha con su nivel de confianza.

Discusiones

Los resultados obtenidos en esta investigación evidencian que la aplicación de técnicas de aprendizaje profundo, específicamente redes neuronales convolucionales combinadas con transferencia de aprendizaje, mejora de manera significativa el reconocimiento automático de emociones faciales. El modelo basado en MobileNetV2 presenta un desempeño superior frente al modelo entrenado desde cero, lo cual confirma que el uso de modelos preentrenados permite aprovechar características visuales previamente aprendidas, reduciendo el sobreajuste y mejorando la capacidad de generalización.

Estos hallazgos concuerdan con estudios previos que destacan la eficacia del transfer learning en tareas de visión por computador, especialmente cuando se

trabaja con datasets de tamaño limitado o con clases desbalanceadas. Asimismo, la integración del modelo en una aplicación web funcional demuestra la viabilidad de aplicar estos enfoques en entornos reales y accesibles para los usuarios.

No obstante, el sistema presenta ciertas limitaciones, particularmente en condiciones de iluminación inadecuada y expresiones faciales ambiguas, lo cual puede afectar la precisión de la predicción. A pesar de ello, el estudio genera una base sólida para futuras investigaciones orientadas a la incorporación de análisis en video en tiempo real, el uso de arquitecturas más avanzadas y la ampliación del conjunto de datos para mejorar el rendimiento del sistema.

Conclusiones

El estudio demuestra que la implementación de redes neuronales convolucionales combinadas con técnicas de transferencia de aprendizaje constituye una solución eficiente y viable para el reconocimiento automático de emociones faciales. El sistema desarrollado alcanzó un nivel de precisión adecuado, evidenciando una mejora significativa frente a modelos entrenados desde cero, y fue integrado exitosamente en una aplicación web funcional e interactiva. Estos resultados validan la aplicabilidad del aprendizaje profundo en proyectos académicos de Ingeniería en Tecnologías de la Información, destacando su potencial para resolver problemas complejos de visión por computador. Asimismo, la investigación sienta las bases para futuros trabajos orientados al análisis de emociones en tiempo real, la incorporación de modelos más avanzados y la ampliación del conjunto de datos, lo que permitiría mejorar la robustez y precisión del sistema. En general, el enfoque propuesto contribuye al desarrollo de soluciones inteligentes con aplicaciones potenciales en áreas como la educación, la salud y la interacción humano-computadora.

Referencias

- Ekman, P., & Friesen, W. V. (2015). Facial Action Coding System: A technique for the measurement of facial movement. Consulting Psychologists Press.
- Goodfellow, I. E. (2015). Challenges in representation learning: A report on three machine learning contests. *Neural Networks*, 64, 59–63.
- Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. *Sensors*, 18(2), 4001.

- Li, S., Deng, W., & Du, J. (2017). Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). AffectNet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1), 18-31.
- Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, 37, 98–125.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*.
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Zhao, G., & Pietikäinen, M. (2007). Dynamic texture recognition using local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (6), 915-9.

Agradecimientos

A la institución por la apertura de publicación

Financiamiento

No aplica

Conflicto de intereses

No existen

El contenido de **ISTE Scientist (ISSN: 2953-6618)**, publicados en este sitio están bajo una licencia de Creative Commons Attribution 4.0 International (CC BY 4.0). Los autores conservan los derechos morales y patrimoniales de sus obras. The contents of this article are under a Creative Commons Attribution 4.0 International (CC BY 4.0) license. The authors retain the moral and patrimonial rights of their works.

