ISBN: 978-628-96613-1-6 ISSN: 2414-6390 Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2025.1.1.1570

Facial recognition of emotions in higher level students during remote classes

Jose Armando Tiznado Ubillus[1]; Luis Antonio Rivera Escriba[2]; Milagros Florencia Mercedes Huamán Martínez[3]; Margarita Betzabé Velásquez Oyola[4]; Fausto Salazar-Fierro[5]; Julio Fernando Arboleda Huaman[6]; César Augusto Atoche Pacherres [7]

[1,3,6]Universidad Nacional Mayor de San Marcos, Lima, Perú, *jose.tiznado@unmsm.edu.pe, milagros.huaman2@unmsm.edu.pe, julio.arboleda@unmsm.edu.pe*

[2]Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, Brasil, *rivera@uenf.br*

[4]Universidad Nacional José Faustino Sánchez Carrión, Huacho, Perú, *mvelasquez@unjfsc.edu.pe*

[5]Universidad Técnica del Norte Facultad de Ingeniería en Ciencias Aplicadas, Ibarra, Ecuador, *fasalazar@utn.edu.ec*

[7]Universidad Nacional de Piura, Escuela de Posgrado. Piura, Perú, *catochep@unp.edu.pe*

Abstract: ***During remote classes, teachers face the challenge of recognizing students' emotions, especially when their cameras are on. To address this problem, a facial detection and recognition system was developed that allowed measuring emotions in real time, providing data that can be statistically analyzed. The objective was to establish and classify the emotions expressed by students during remote classes, which allows detecting their real-time participation in the virtual environment. To obtain the best predictive model for detecting student participation in real time, different models were adjusted. Various reference data sets were used to measure the performance and accuracy of the proposed system. The results showed that the proposed system achieves accuracy on different reference sets and on the proposed data set itself. The real-time predictive classification model, modelFEC6.h5, outperforms the others with an accuracy of 94.62% for facial emotion classification in real-time virtual learning scenarios.***

Keywords: ***Emotion recognition, real-time facial detection, remote classes, virtual environment interaction***

Reconocimiento facial de emociones en estudiantes de nivel superior durante clases remotas

Jose Armando Tiznado Ubillus[1]; Luis Antonio Rivera Escriba[2]; Milagros Florencia Mercedes Huamán Martínez[3]; Margarita Betzabé Velásquez Oyola[4]; Fausto Salazar-Fierro[5]; Julio Fernando Arboleda Huaman[6]; César Augusto Atoche Pacherres [7]

[1,3,6]Universidad Nacional Mayor de San Marcos, Lima, Perú, *jose.tiznado@unmsm.edu.pe, milagros.huaman2@unmsm.edu.pe, julio.arboleda@unmsm.edu.pe*

[2]Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, Brasil, *rivera@uenf.br*

[4]Universidad Nacional José Faustino Sánchez Carrión, Huacho, Perú, *mvelasquez@unjfsc.edu.pe*

[5]Universidad Técnica del Norte Facultad de Ingeniería en Ciencias Aplicadas, Ibarra, Ecuador, *fasalazar@utn.edu.ec*

[7]Universidad Nacional de Piura, Escuela de Posgrado. Piura, Perú, *catochep@unp.edu.pe*

Resumen: ***Durante las clases remotas, los docentes enfrentan el reto de reconocer las emociones de los estudiantes, especialmente cuando sus cámaras están encendidas. Para abordar este problema, se desarrolló un sistema de detección y reconocimiento facial que permitió medir las emociones en tiempo real, proporcionando datos que pueden analizarse estadísticamente. El objetivo fue establecer y clasificar las emociones expresadas por los estudiantes durante el dictado de clases remotas, que permitan detectar su participación en tiempo real en el entorno virtual. A fin de obtener el mejor modelo predictivo en la detección de la participación del estudiante en tiempo real, se ajustaron diferentes modelos. Se utilizaron diversos conjuntos de datos de referencia al medir el rendimiento y la precisión del sistema propuesto. Los resultados mostraron que el sistema alcanza una precisión en diferentes conjuntos de referencia y en el propio conjunto de datos propuestos. El modelo de clasificación predictiva en tiempo real modelFEC6.h5, supera a los demás con una precisión del 94.62% para la clasificación de emociones faciales en el escenario del aprendizaje virtual en tiempo real.***

Palabras Clave: Reconocimiento de emociones, detección facial en tiempo real, clases remotas, interacción en entorno virtual.

I. INTRODUCCIÓN

Los rostros humanos tienen la notable capacidad de expresar emociones a través de contracciones y relajaciones musculares [1]. Así lo demuestra la investigación [2] que originalmente estudió los movimientos y encontró que las expresiones musculares están relacionadas con los gestos. Al respecto, otro estudio [3] afirma que, las expresiones de los rostros pueden ser manifestadas a través de diferentes tipos de emociones, las que se sintetizan en siete categorías universales: felicidad, tristeza, ira, miedo, sorpresa, desprecio y asco. En este camino de avance de la tecnología, algunos autores [4, 5] convergen en que el aprendizaje profundo (DL) forma parte de la inteligencia artificial (IA) y del aprendizaje automático (ML). Es así como la IA, se puede relacionar con la capacidad humana de clasificar gestos emocionales mediante redes neuronales artificiales inspiradas en el cerebro [5]. Los sistemas modernos actuales aplican reconocimiento de emociones a partir de expresiones faciales (FER), construyendo típicamente, combinando algoritmos, modelos de aprendizaje y datasets de imágenes o videos. No obstante, a pesar de los avances; estos aún no llegan a la precisión exacta en el reconocimiento de las emociones desde diferentes ángulos requiriendo continuar en el camino del entrenamiento [6, 7]. En este marco, se encontró que, al analizar un número mayor de datos, los algoritmos básicos permanecen sin cambios, pero los pesos y sesgos internos del modelo, que son clave para elegir una respuesta específica, se ajustan continuamente [4].

En este contexto de aprendizaje adaptativo, las emociones de los estudiantes durante las clases virtuales juegan un papel importante en la adquisición del aprendizaje, para lo cual el docente deberá apoyarse a través de una metodología realmente innovadora y didáctica que no se limite meramente a la impartición de clases a través de una cámara [8]. De esta manera, la implementación de la tecnología y el uso de medios de apoyo pedagógico enriquecen el ambiente de aprendizaje y motivan a los estudiantes [9]. Por lo tanto, los profesionales comprometidos con el área de la Educación deben ser conscientes del importante papel que desempeñan para influir en la motivación del alumno al diseñar actividades de aprendizaje [10].

Así, un estudio [11], proporciona evidencia detallada sobre el registro de emociones negativas en sesiones de clases virtuales en el contexto de la pandemia de COVID-19, mientras que, en otro estudio [12] referente a la educación en línea aplicada en el ámbito de la rehabilitación médica, advierte que, las emociones negativas, como la ansiedad y la depresión, pueden afectar la adherencia y el aprovechamiento de los programas educativos diseñados para pacientes. Otra investigación [17] deja traslucir uno de los desafíos psicológicos que ocurre en el aprendizaje en línea, la presencia de emociones negativas como la ansiedad y el aburrimiento en estudiantes chinos que están aprendiendo un segundo idioma. Otro estudio [13], describe un caso en Grecia, pero con estudiantes de nivel primaria mostrando conducta de comportamientos negativos en sus emociones asociados a factores psicológicos, educativos y tecnológicos afectando también a los docentes y padres involucrados. La misma situación ocurre respecto a la falta de atención de la expresión emocional en el aprendizaje dentro del curso de matemática, mostrando como resultado emociones negativas, generando ansiedad en el curso

afectando significativamente el compromiso y la compresión de los estudiantes en el entorno de e-learning [14]. Un grupo de educadores en hospitalidad de enseñanza remota de emergencia (ERT) expresó emociones negativas, como ansiedad, estrés y frustración, debido a problemas técnicos de conectividad, fallas técnicas y sistemas incompatibles en la en la modalidad remota [15]. A continuación, se presenta una lista de emociones negativas asociadas con la educación en línea en distintos contextos y países (Tabla 1).

Tabla 1: Emociones negativas asociadas con la educación

Autor	Modalidad	País	Emociones
[3]	Clases Virtuales	China	Enojo, Miedo, Disgusto, Tristeza
[10]	Educación en línea	China	Ansiedad, Depresión.
[12]	aprendizaje en línea	China	Ansiedad, Aburrimiento.
[13]	Enseñanza en línea	Grecia	Ansiedad, Tristeza, Aburrimiento, Frustración Aislamiento, Falta de interés
[15]	Remota	Asia-Pacífico	Frustración, Estrés, Ansiedad, Agotamiento, Apatía, Desesperanza, Culpabilidad, Desamparo

Fuente: Elaboración propia.

Al respecto, en un estudio [6] se describen las emociones positivas involucradas, como la felicidad y la neutralidad. Otra investigación [15] analiza las experiencias emocionales de los educadores en hospitalidad durante la enseñanza remota en épocas de emergencia (ERT), destacando transiciones emocionales positivas, entre las que se encuentran la alegría, la sorpresa y la satisfacción. Además, de la investigación [16] que relaciona la alegría como una emoción positiva. A estos hallazgos se agrega que, las emociones positivas, como el disfrute, están vinculadas al aprendizaje en línea, fomentando la motivación y el compromiso de los estudiantes, lo que mejora su rendimiento y bienestar emocional en contextos educativos [17].

En este camino por alcanzar la precisión, se fueron identificando ciertas lagunas del conocimiento, es así que uno de los vacíos encontrados [6], describe que incorporar datos recolectados en situaciones más diversas, como diferentes posturas o condiciones de iluminación, mejora el rendimiento del modelo de detección de emociones para incrementar la precisión especialmente en emociones negativas que presentaron mayor variabilidad.

Es así como, al revisar el estado de arte, se encontró un antecedente importante [18] que muestra el análisis exhaustivo sobre las expresiones faciales. La investigación ha sido una referencia fundamental en el desarrollo de sistemas computacionales diseñados a fin de reconocer emociones a través de diversas modalidades de entrada; como el rostro, la voz y las expresiones faciales. Sin embargo, la mayoría de los investigadores han adoptado técnicas de Deep Learning para el reconocimiento de emociones faciales [16], presentando un enfoque basado en el modelo RACN, que combina redes neuronales convolucionales (CNN) con características estructurales adaptativas, logrando mejorar la precisión en la clasificación de emociones. En el estudio, evaluaron el rendimiento del modelo en los conjuntos de datos CK+, FER2013 y RAF-DB, alcanzando tasas de precisión de 98.9%, 74% y 88%, respectivamente, lo que demuestra la eficacia del método en escenarios controlados y reales. De manera similar, otra investigación [25] aplicó técnicas CNN, complementando su enfoque e incorporando a su vez redes completamente convolucionales (FCN) para mejorar la detección de emociones. Utilizando el conjunto de datos RAFD con seis expresiones y la etiqueta happy, lograron una precisión del 96% con FCN y del 93% al emplear CNN sobre el mismo conjunto de datos. Otros investigadores [23], implementaron modelos avanzados de CNN, como ResNet-50, VGG-19 e Inception-V3, evaluándolos en datasets públicos (FER2013, CK+, RAF-DB) y en un dataset propio. Los resultados obtenidos variaron entre 73.40% y 92.32%, dependiendo del modelo y la combinación de datasets utilizados. En particular, ResNet-50 mostró el mejor desempeño, alcanzando una precisión del 92.32% al entrenarse con la combinación de FER2013, CK+, RAF-DB y un dataset propio, lo que confirma la efectividad de utilizar múltiples fuentes de datos para mejorar la generalización del modelo. Además, otro estudio [19] adoptó un enfoque innovador al aplicar Inteligencia Artificial Explicable (XAI) analizando la relación entre el arte laqueado y las emociones, combinando técnicas de Deep Learning con Machine Learning, utilizando CaffeNet para la extracción de características y Random Forest para la clasificación, junto con métodos de explicabilidad como LIME y SHAP. Los resultados obtenidos muestran una precisión del 92%, superando a otros modelos basados en CNN (88%) y al propio Random Forest (86%), lo que resalta la eficacia del modelo propuesto y su capacidad para interpretar las decisiones del sistema.

Otro estudio [20] propone la necesidad de desarrollar un sistema para reconocer las emociones faciales de los estudiantes por parte del docente durante el dictado de las clases remotas y estimar dicha sesión.
La presente investigación tuvo como objetivo general establecer y clasificar las emociones expresadas por los estudiantes durante el dictado de clases remotas. Los objetivos específicos que los guiaron fueron:

- Elaborar un modelo IA y entrenarlo con datasets publica para estimar buena precisión en el reconocimiento de las emociones faciales.
- Comprobar la eficacia del modelo IA con datasets publica, desarrollando algoritmos adaptados para optimizar su aplicabilidad y precisión al contexto educativo de Perú y Ecuador.

II. MÉTODO

El propósito de alcanzar el objetivo establecido llevó a utilizar técnicas basadas en redes neuronales convolucionales (CNN) y algoritmos de Deep Learning aplicados a datases públicos. Estas herramientas permitieron entrenar un modelo diseñado para el reconocimiento de emociones faciales a partir de imágenes procesadas, facilitando así su caracterización y análisis a través de redes neuronales.

El proyecto se implementó con tres etapas claves: la primera, llevó a cabo la limpieza y el pre procesamiento de los datasets públicos seleccionados, asegurando que los datos

fueran adecuados y pertinentes para el proceso de entrenamiento; la segunda, procedió al entrenamiento del modelo de inteligencia artificial utilizando los datos previamente preparados. Por último, la tercera, desarrolló un software en Python que integró el modelo entrenado, permitiendo el reconocimiento de emociones faciales en tiempo real a partir de las interacciones de los estudiantes.

El estudio tuvo en cuenta el marco ético aplicado al reconocimiento facial basado en el uso de los datos masivos que incluyó: a) el consentimiento para el uso de los datos personales; b) la validez y fiabilidad de las aplicaciones tecnológicas; c) los posibles impactos adversos y; d) la participación de los distintos colectivos de la comunidad en la aplicación de tecnología basada en datos [27].

Este enfoque metodológico buscó garantizar un sistema eficiente y confiable, promoviendo avances en la identificación automatizada de estados emocionales en contextos educativos. La Figura 1 muestra el desarrollo general del proyecto.

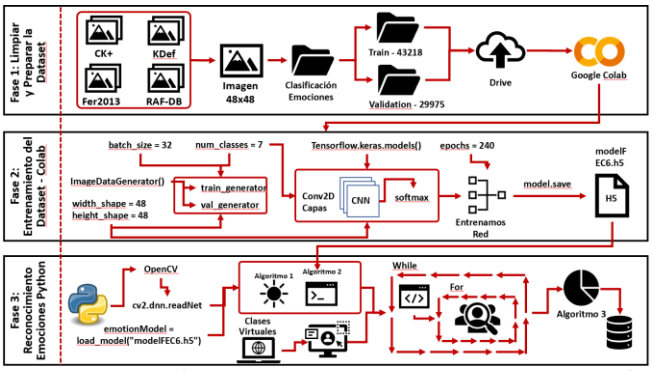


Figura 1: Descripción general del desarrollo del proyecto. Elaboración Propia.

Se llevaron a cabo múltiples pruebas iniciales en las que, la precisión obtenida no cumplió con los niveles esperados. Este resultado hizo necesario experimentar con diferentes data sets, aplicando técnicas avanzadas de Deep Learning con el objetivo de optimizar la precisión del modelo de inteligencia artificial. Tras un proceso iterativo de ajustes y evaluaciones, se logró alcanzar una tasa de confiabilidad adecuada, lo que permitió la implementación del modelo para la identificación efectiva de emociones en estudiantes. A continuación, se describen las fases de implementación y las pruebas realizadas durante el desarrollo del proyecto.

A. Limpieza y preparación de la Data set

En esta fase, se llevó a cabo la preparación y selección de los data set públicos (FER2013, CK+, KDEF, RAF-DB) tal como se muestra en la tabla 2. Inicialmente, se trabajó con un único data set, sobre el cual se ejecutaron pruebas para evaluar su efectividad. Posteriormente, se realizó la combinación de múltiples data sets públicos con el fin de mejorar la representatividad de los datos y alcanzar la precisión deseada. Este proceso de la Figura 2 fue fundamental para garantizar la calidad y consistencia de los datos utilizados en el entrenamiento del modelo.

Tabla 2: Lista de conjunto de Dataset de rostro públicos

Dataset	Acceso	Correo	Oficial
FER2013	Publico	No requiere	https://www.kaggle.com/datasets/msambare/fer2013
CK+	Publico	No requiere	https://www.kaggle.com/datasets/shuvoalok/ck-dataset
KDEF	Cerrado	Autorizado, citado	https://www.kaggle.com/datasets/tom99763/testtt https://kdef.se/faq/using-and-publishing-kdef-and-akdef
RAF-DB	Cerrado	Falta respuesta	https://www.kaggle.com/datasets/shuvoalok/raf-db-dataset/data http://www.whdeng.cn/raf/model1.html#dataset

La datasets requirió de autorización KDEF, Raf-DB; esta base de datos constituida por más de 300 mil imágenes faciales extraídas de internet clasificada en siete categorías: sorpresa, miedo, asco, ira, tristeza, felicidad y neutralidad [22, 21].

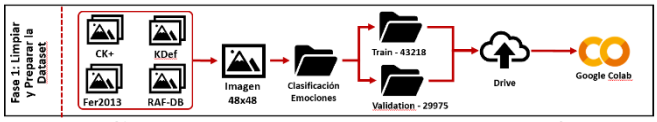


Figura 2: Gráfico de limpieza y selecciona de Dataset. Elaboración propia.

GoogleColab:https://colab.research.google.com/drive/1m46u2KE6dLj2xQNitavhekqoeFfz3FrZ?usp=sharing

1) Preparación del entorno y extracción del dataset: Se instaló la biblioteca rarfile para manejar archivos comprimidos en formato "rar", definiendo un procedimiento para extraer el contenido del archivo comprimido en una carpeta específica, asegurando que la estructura del dataset original estuviera disponible para su procesamiento.

2) Redimensionamiento y conversión a escala de grises: Se desarrollaron dos funciones principales para procesar las imágenes; a) Redimensionamiento individual: Se ajustaron las dimensiones de cada imagen a un tamaño fijo de 48x48 píxeles. b) Conversión a escala de grises: Cada imagen fue convertida al formato monocromático para simplificar su representación y reducir la complejidad de los datos. El procesamiento se aplicó de manera iterativa a todas las imágenes en las carpetas del dataset, preservando la estructura original de las carpetas en el directorio de salida.

3) Almacenamiento de las imágenes procesadas: Las imágenes procesadas fueron almacenadas en una nueva carpeta, donde se replicó la organización jerárquica de los datos originales para facilitar el acceso y análisis posterior.

4) Compresión y exportación del dataset procesado: La carpeta que contenía las imágenes procesadas fue comprimida en un archivo ZIP, garantizando que no hubiera conflictos de nombres en caso de archivos previos. Finalmente, el archivo comprimido fue preparado para descarga mediante un enlace generado automáticamente.

B. Entrenamiento del Dataset - Colab

En esta sesión se procedió a entrenar la IA empleando Google Colab de Python que utiliza TensorFlow y Keras para construir, entrenar y evaluar un modelo de red neuronal convolucional (CNN) para el reconocimiento de expresiones faciales. A continuación, se describe paso a paso lo que sucede en la Figura 3

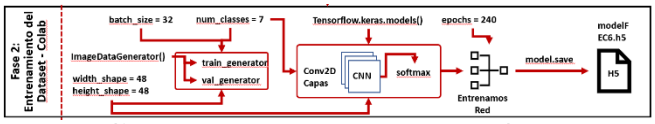


Figura 3: Gráfico de Entrenamiento del Dataset. Elaboración propia

GoogleColab:https://colab.research.google.com/drive/1AASoaE5cI2ETg-xSkf5OVt_XLmLwuWuI?usp=sharing

1) Importación de librerías y montaje de Google Drive: Se importan las librerías necesarias de TensorFlow, Keras y otras bibliotecas de Python para construir y entrenar el modelo. También se monta la unidad de Google Drive para acceder a los datos.

2) Definición de rutas y parámetros: Se definen las rutas de los directorios de entrenamiento y validación, así como los parámetros del modelo, como el tamaño de la imagen de entrada (48x48 píxeles), el número de clases (7 emociones), el número de épocas (240) y el tamaño del lote (32). El parámetro epochs = 240 especifica el número de veces que se iterará sobre el conjunto de datos completo durante el entrenamiento.

3) Preparación de los datos: El conjunto de datos utilizado proviene de cuatro fuentes diferentes: CK+, KDef, Fer2013 y Raf-Db. Estas fuentes contienen imágenes de expresiones faciales etiquetadas con diferentes emociones. El conjunto de datos completo se encuentra comprimido en un archivo llamado "images3.5.rar". Primero, se descomprime el archivo "images3.5.rar" en el directorio actual utilizando el comando: [!unrar x images3.5.rar /content]

Esto extrae el contenido del archivo comprimido en una carpeta llamada "images3.5". Dentro de esta carpeta, se encuentran dos subcarpetas: "train" y "validation".
La carpeta "train" contiene 43,218 imágenes de entrenamiento, mientras que la carpeta "validation" contiene 29,975 imágenes de validación. Ambas carpetas tienen subcarpetas etiquetadas con los nombres de las emociones: 'angry', 'disgust', 'fear', 'happy', 'neutral', 'sad' y 'surprise'.
Se utilizan los generadores de datos de Keras (ImageDataGenerator) para cargar y preprocesar las imágenes de entrenamiento y validación. Estos generadores aplican técnicas de aumento de datos, como rotaciones, cambios de escala, etc., para mejorar la capacidad de generalización del modelo. Las imágenes se convierten a escala de grises y se redimensionan a 48x48 píxeles.

4) Construcción del modelo CNN: La Figura 4 muestra la arquitectura secuencial de una red neuronal convolucional (CNN) diseñada para tareas de clasificación de imágenes. Esta red se construyó utilizando las siguientes componentes principales:

(a) Instancia de Sequential: Permite construir el modelo capa por capa de forma ordenada.

(b) Capas Convolucionales (Conv2D): Cada capa aplica una operación de convolución definida

$$(I * K)(x, y) = \sum_m \sum_n I(m, n).K(x - m, y - n)$$

donde ***I*** es la imagen de entrada, K es el kernel, x y y son las coordenadas de salida, y m y n son los índices del kernel. Los tamaños de kernel y el número de filtros varían según la capa para extraer características de diferentes niveles.

(c) Normalización por Lotes y ReLU: La normalización por lotes (BatchNormalization) regula las activaciones intermedias para acelerar el aprendizaje.
La función de activación ReLU *(f(x)=max(0,x))* introduce no linealidad en el modelo.

d) MaxPooling y Dropout:
MaxPooling: Reduce la dimensionalidad seleccionando el valor máximo en regiones locales.
Dropout: Previene el sobreajuste desactivando aleatoriamente algunas neuronas durante el entrenamiento.

(e) Aplanamiento y Capas Densas: Después de las capas convolucionales, la salida se aplana (Flatten) para pasar a las capas densas.
Se incluyen dos capas densas con 128 y 256 neuronas, cada una normalizada y activada mediante ReLU, con abandono aplicado para regularización.

(f) Capa de Salida: Contiene tantas neuronas como clases (en este caso, 7 emociones) y utiliza la activación Softmax para calcular probabilidades:

$$softmax(x_i) = \frac{exp(x_i)}{\sum_j exp(x_j)}$$

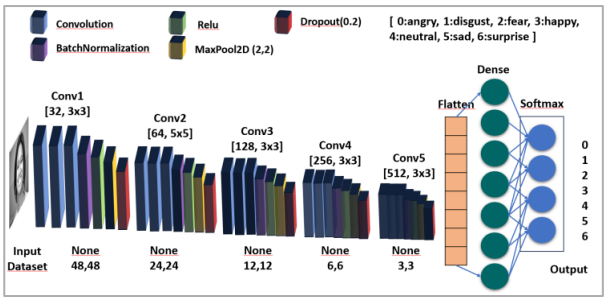


Figura 4: Diagrama de Arquitectura CNN. Elaboración propia

5) Resultados Experimentales: Configuración del Modelo: El modelo fue construido utilizando la función de pérdida categorical_crossentropy, el optimizador Adam, y la métrica accuracy. La función categorical_crossentropy se seleccionó por ser estándar para problemas de clasificación multiclase, mientras que el optimizador Adam combina las ventajas de RMSProp y SGD, mejorando la convergencia en datasets complejos. El modelo se entrenó utilizando generadores de datos de entrenamiento y validación con técnicas de aumento de datos (data augmentation) para incrementar la diversidad del dataset. Se configuró TensorBoard para monitorear las métricas de entrenamiento y detectar posibles problemas, como el sobreajuste.

Resultados del Entrenamiento: Se realizaron tres experimentos, variando el número de imágenes, los datasets utilizados y las épocas de entrenamiento. Los resultados se resumen en la siguiente Tabla 3.

Tabla 3: Resultados Comparativos de Entrenamientos del Modelo

Train	Datasets	Img	Epoch	Precision Prom.	F1-Score Prom.	Accuracy(%)	Mejor Modelo
Train 1	[FER2013,CK+,KDef]	15350	150	0.7267	0.7272	0.7347	No
Train 2	[FER2013,CK+,KDef]	18191	240	0.9089	0.9085	0.9099	No
Train 3	**[FER2013,CK+, KDef, Raf-DB]**	**29755**	**240**	**0.9464**	**0.9458**	**0.9462**	**Si**

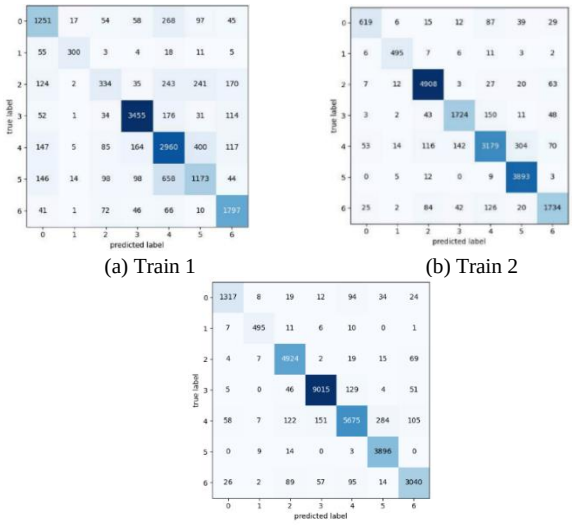
Fuente: Elaboración propia.

6) Análisis de las Matrices de Confusión: Las matrices de confusión (Figura 5) permiten analizar en detalle el desempeño del modelo en cada clase. Las filas representan las etiquetas reales, mientras que las columnas muestran las etiquetas predichas.

(a) Train 1: Observaciones: La matriz muestra una notable confusión entre las clases 2 y 6, lo que podría atribuirse a similitudes en sus características visuales. El modelo alcanzó un accuracy del 73.47%.

(b) Train 2: Observaciones: Con un dataset ampliado (18,191 imágenes) y 240 épocas, se observó una mejora significativa en la precisión, con la mayoría de las predicciones concentradas en la diagonal principal. Se alcanzó un accuracy del 90.99%.

(c) Train 3: Observaciones: Este modelo, entrenado con 299,755 imágenes y un dataset más diverso (incluyendo Raf-DB), presentó el mejor desempeño. La matriz refleja una mínima confusión entre clases, con una mayor densidad de valores en la diagonal principal. El accuracy alcanzó el 94.62%.



(a) Train 1 (b) Train 2

(c) Train 3

Figura 5: Train1(a), Train2(b) y Train3(c) son Matriz de Confusión

Desempeño por Clase: El análisis detallado por clase se muestra en la siguiente Tabla 4

Tabla 4: Métrica de Clasificación por Etiqueta

Clase	Precisión	Recall	F1-Score	Soporte
0	0.9294	0.8733	0.9005	1508
1	0.9375	0.934	0.9357	530
2	0.9424	0.977	0.9594	5040
3	0.9753	0.9746	0.975	9250
4	0.9419	0.8864	0.9133	6402
5	0.9174	0.9934	0.9538	3922
6	0.924	0.9148	0.9194	3323
accuracy			0.9462	29975
macro avg	0.9383	0.9362	0.9367	29975
weighted avg	0.9464	0.9462	0.9458	29975

7) Visualización de las Matrices de Confusión: Figura 5: Matrices de confusión para (a) Train 1, (b) Train 2, y (c) Train 3. La diagonal principal refleja las predicciones correctas, mientras que los valores fuera de la diagonal representan errores de predicción.

Las matrices correspondientes a cada configuración de entrenamiento se presentan en la Figura 5. Estas matrices muestran la relación entre las predicciones y las etiquetas reales, con la mayoría de los valores concentrados en la diagonal principal.

Las métricas estándar de clasificación, como precisión, recall y F1-score, se calcularon para cada clase utilizando las siguientes fórmulas:

$$Precisión = \frac{TP}{TP + FP}$$

$$Recall\ (Sensibilidad) = \frac{TP}{TP + FN}$$

$$F1 - Score = 2 \cdot \frac{Precisión\ .\ Recall}{Precisión + Recall}$$

donde TP representa los verdaderos positivos, FP los falsos positivos, y FN los falsos negativos.

El modelo entrenado en Train 3 presentó el mejor desempeño, con un accuracy del 94.62%, mostrando una clasificación precisa de emociones faciales en un dataset diverso. Las áreas de confusión identificadas, como entre las clases 2 y 6, representan oportunidades para mejorar la generalización del modelo.

8) Guiando el modelo entrenado: El modelo CNN fue diseñado para clasificar emociones faciales en 7 categorías, utilizando los datasets CK+, KDef, Fer2013 y Raf-DB, con un total de 299,755 imágenes. Entrenado durante 240 épocas, alcanzó un accuracy del 94.62%, con una precisión promedio ponderada de 94.64% y un F1-score promedio de 94.58%. Los resultados detallados por clase se presentan en la (Tabla 4), destacando el excelente desempeño en la clase 3 (F1-score de 0.975) y áreas de mejora como la confusión en las clases 2 y 6. El modelo fue guardado en formato .h5, facilitando su reutilización y despliegue en aplicaciones prácticas. La diversidad del dataset y el optimizador Adam fueron claves para lograr estos resultados robustos.

C. Reconocimiento de emociones en python

Se presenta una visión general Figura 6 del flujo de trabajo diseñado para el reconocimiento de emociones faciales en un entorno de clases remotas utilizando Python, OpenCV, y un modelo preentrenado en formato .h5.

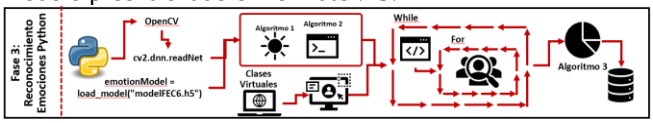


Figura 6: Fase 3 de reconocimiento de Emociones en Python

El sistema está compuesto por los siguientes componentes principales:

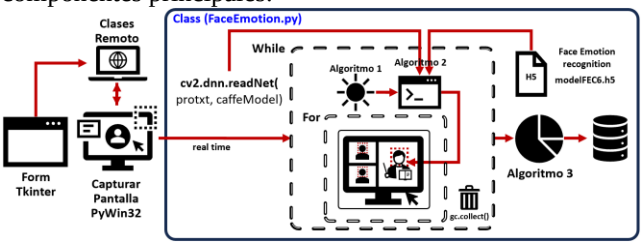


Figura 7: Arquitectura del Sistema FaceEmotion - RealTime

1) Inicio del Sistema: Formulario Tkinter: Proporciona una interfaz gráfica que permite al usuario iniciar el sistema y configurar los parámetros iniciales. Desde aquí, el usuario puede seleccionar la clase virtual que desea analizar como se muestra en la Figura 8.

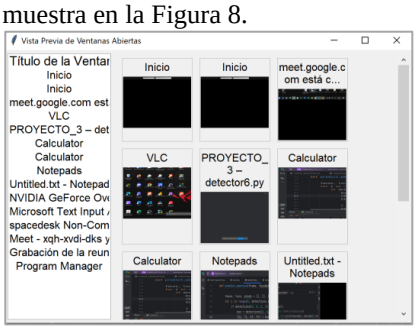


Figura 8: Formulario de lista de ventanas activas

Captura de Pantalla PyWin32: Utiliza la biblioteca PyWin32 para capturar en tiempo real las imágenes de la pantalla de las clases remotas. Estas imágenes servirán como entrada para los algoritmos de análisis.

2) Procesamiento en Tiempo Real: El procesamiento se gestiona en un bucle principal (While) dentro de la clase FaceEmotion.py, en el que se analizan las imágenes capturadas siguiendo estos pasos:

Algoritmo 1 - Mejora de la Imagen: Se mejora la iluminación y el contraste de las imágenes capturadas para garantizar una detección precisa de los rostros. Este preprocesamiento es crucial para optimizar la calidad de las entradas al modelo. El algoritmo 1 Preprocess_Face optimiza imágenes faciales capturadas en condiciones de iluminación extremas para su análisis posterior. Convierte la imagen a escala de grises, aplica corrección Gamma para ajustar el brillo y normaliza el histograma para mejorar el contraste global. En el espacio de color LAB, utiliza CLAHE para ajustar la luminancia localmente. Luego, reduce el ruido con un filtro bilateral que preserva los bordes importantes, redimensiona la imagen a 48x48 píxeles y la convierte en un arreglo numérico listo para modelos de análisis. Este proceso garantiza imágenes de alta calidad, mejorando la precisión en tareas como reconocimiento facial o análisis emocional.

Algoritmo 1	Preprocess_Face - Procesamiento de imágenes faciales optimizado para la iluminación
Entrada:	Una imagen **facial** capturada en condiciones extremas de iluminación.
Salida:	Una **imagen preprocesada** y ajustada lista para análisis posterior.
Empezar	
	1. Convertir la imagen **facial** a escala de grises. 2. Aplicar corrección Gamma: a. Definir el valor de Gamma (γ=1.2). b. Calcular el inverso de Gamma (invGamma = 1/γ). c. Construir una tabla de mapeo de valores de intensidad. d. Aplicar la tabla usando la función LUT. 3. Normalizar el histograma de la imagen globalmente usando equalizeHist. 4. Ajustar luminancia en el espacio de color LAB: a. Convertir la imagen al espacio de color LAB. b. Dividir la imagen en canales L, A, B. c. Aplicar CLAHE en el canal L co: clipLimit = 3.0 tileGridSize = (8, 8). d. Reconstruir la imagen combinando los canales ajustados. e. Convertir nuevamente al espacio de grises. 5. Suavizar el ruido aplicando un filtro bilateral: Radio del filtro: 9. Valores Sigma para intensidad y espacio: 75, 75. 6. Redimensionar la imagen a 48x48 píxeles. 7. Convertir la imagen a un arreglo de flotantes sin normalización. 8. Expandir las dimensiones del arreglo para adaptarlo al modelo. 9. Devolver la **imagen preprocesada**.
Fin	

Algoritmo 2 - Reconocimiento de Emociones: Se utiliza un modelo preentrenado (modelFECC6.h5) cargado mediante cv2.dnn.readNet(protxt, caffeModel). Este modelo identifica las emociones faciales en tiempo real, clasificándolas en categorías como alegría, tristeza, enojo, sorpresa, etc. El algoritmo 2 Predict_Emotion detecta rostros y predice emociones asociadas. Aumenta la resolución de la imagen, utiliza faceNet para identificar rostros y filtra las detecciones según un umbral de confianza. Recorta y preprocesa los rostros con Preprocess_Face, almacena sus ubicaciones y utiliza emotionModel para predecir emociones. Devuelve las ubicaciones de los rostros y las predicciones emocionales, permitiendo un análisis eficiente y preciso en tiempo real.

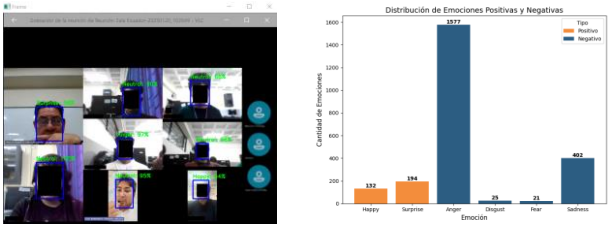
Algoritmo 2:	Predict_Emotion - Predicción de emociones basada en detección de rostros
Entrada:	Una imagen (**frame**) capturada en tiempo real, modelo de detección facial (**faceNet**) y el modelo de predicción de emociones (**emotionModel**).
Salida:	Ubicaciones de rostros detectados (**locs**) y predicciones de emociones asociadas (**preds**).
Procedimiento:	1. Aumentar la resolución de la imagen para mejorar la detección de rostros pequeños: Crear un blob a partir de la imagen usando cv2.dnn.blobFromImage. Asignar el blob al modelo de detección facial faceNet. Ejecutar la red neuronal para obtener detecciones (faceNet.forward). 2. Inicializar listas vacías para rostros (faces), ubicaciones (locs) y predicciones (preds). 3. Iterar sobre cada detección obtenida usando un bucle for: a. **Para cada detección i:** Verificar si la confianza de la detección supera el umbral (e.g., 0.3) usando una instrucción if: **Si** la confianza detections[0, 0, i, 2] > 0.3, proceder: 1. Calcular las coordenadas de la caja delimitadora ajustándolas al tamaño de la imagen original: Usar las coordenadas normalizadas proporcionadas por detections[0, 0, i, 3:7]. Multiplicarlas por las dimensiones de la imagen para obtener valores en píxeles. Asegurarse de que las coordenadas iniciales (Xi, Yi) sean mayores o iguales a 0. 2. Recortar la región facial de la imagen usando las coordenadas obtenidas (face = frame[Yi:Yf, Xi:Xf]). 3. Preprocesar la imagen facial recortada llamando a la función **Algoritmo1 - Preprocess_Face(faces)** para optimizar la iluminación. 4. Agregar la imagen preprocesada a la lista faces. 5. Almacenar las coordenadas de la región facial en la lista locs. 6. Realizar la predicción de emociones: Pasar la imagen preprocesada al modelo de emociones (emotionModel). Obtener la predicción (pred = emotionModel.predict(face)[0]). Almacenar la predicción en la lista preds. 4. Repetir el proceso para todas las detecciones válidas en el bucle for. 5. Devolver las listas de ubicaciones (**locs**) y las predicciones de emociones (**preds**).
Fin del algoritmo	

Algoritmo 3 - Generación de Resultados: Los datos obtenidos se analizan para generar gráficos que clasifican las emociones detectadas en positivas y negativas. Los resultados se almacenan en una base de datos para análisis futuros. El algoritmo 3 Gráfico_Emotion genera un gráfico de barras agrupadas para visualizar emociones positivas (Happy, Surprise) y negativas (Anger, Disgust, Fear, Sadness). Utiliza una lista de datos para crear un DataFrame, filtra las emociones relevantes y las clasifica como positivas o negativas. Luego, agrupa los datos y cuenta las ocurrencias por emoción. Configura un gráfico de barras con colores distintivos (naranja para positivas y azul para negativas), etiqueta los ejes y agrega un título y leyenda. Finalmente, muestra el gráfico, facilitando el análisis visual de la distribución de emociones detectadas.

Algoritmo 3	Grafico_Emotion - Generar gráfico estadístico de emociones
Entrada:	Una lista de datos data que contiene emociones detectadas con sus valores porcentuales.
Salida:	Un gráfico de barras agrupadas que representa las emociones positivas y negativas.
Empezar	
	1. Definir las categorías de emociones: Emociones positivas: Happy, Surprise. Emociones negativas: Anger, Disgust, Fear, Sadness. 2. Crear un DataFrame df usando los datos proporcionados en data. 3. Filtrar las emociones para incluir únicamente las categorías positivas y negativas. 4. Clasificar las emociones: Para cada emoción C en el DataFrame df: Si C pertenece a emociones positivas, asignar el tipo "Positivo". Si C pertenece a emociones negativas, asignar el tipo "Negativo". 5. Agrupar los datos por emoción y tipo, y contar la cantidad de ocurrencias: **Cuenta total por emoción =** $\sum_{i=1}^{n} V_i$ 6. Configurar el gráfico de barras: a. Crear un gráfico de barras agrupadas con dos grupos: Emociones Positivas en color naranja (#F38D3C). Emociones Negativas en color azul (#2C5D82). b. Asignar las etiquetas de los ejes: Eje X: "Emoción". Eje Y: "Cantidad de Emociones". 7. Agregar leyenda y título: Título: "Distribución de Emociones Positivas y Negativas". Leyenda: "Positivo" y "Negativo". 8. Mostrar el gráfico generado.
Fin	

3) Optimización de Recursos: Eliminación de Datos Temporales (gc.collect()): Al final de cada iteración del bucle, se eliminan los datos temporales generados durante el procesamiento para optimizar el uso de memoria y garantizar la estabilidad del sistema.

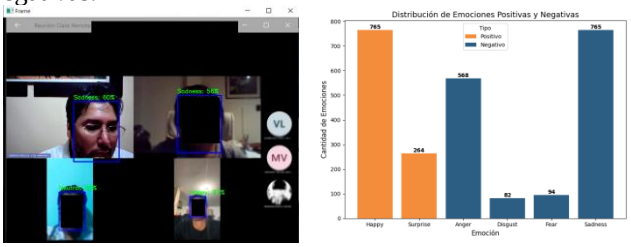
4) Gráficos de Emociones: Los resultados del análisis se representan gráficamente, lo que facilita la interpretación de los estados emocionales de los participantes. Almacenamiento en Base de Datos en un Excel: Los datos procesados se guardan para su consulta posterior, permitiendo evaluaciones continuas de las emociones en el contexto de las clases remotas en Ecuador.



(a) Software (b) Gráfico de Barra Ecuador

Figura 9: Software(a), Grafico Barra Ecuador(b) para reconocer las emociones básicas clasificada con clasificadas en emociones positivas y negativas

En la Figura 9, se describe que el tiempo transcurrido sobre la reproducción sobre el video duro con un total de 23 minutos con 40 segundo que se capturaron todas las emociones en forma general con un aproximado de 8 estudiantes que activaron sus webcams, que al final muestra Figura 9 su clasificación en emociones positivas y negativas.



(a) Software (b) Gráfico de Barra Ecuador

Figura 10: Software(a), Grafico Barra Perú(b) para reconocer las emociones básicas clasificada con clasificadas en emociones positivas y negativas

En la Figura 10, se describe que el tiempo transcurrido sobre la reproducción sobre el video duro con un total de 24 minutos con 51 segundo que se capturaron todas las emociones y se guardó en un Excel de forma general con un aproximado de 5 estudiantes que activaron sus webcams, que al final muestra Figura 10 su clasificación en emociones positivas y negativas.

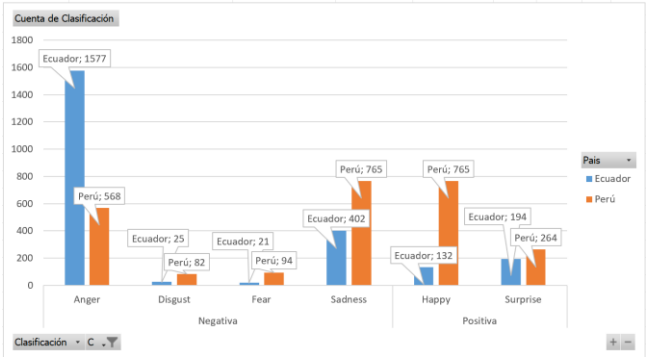


Figura 11: Estadística comparativa de Emociones Entre Perú y Ecuador

En la Figura 11 se presenta una comparación de las emociones más frecuentes según su clasificación en positivas y negativas, diferenciadas entre Perú y Ecuador. Este análisis tiene como propósito identificar patrones emocionales predominantes en cada país, con el objetivo de evaluar cómo las emociones positivas pueden influir en mejoras potenciales dentro del contexto educativo, fomentando ambientes de aprendizaje más efectivos y adaptados a las necesidades emocionales de los estudiantes.

III. RESULTADOS

1. Elaborar un modelo IA y entrenarlo con datasets público para estimar buena precisión en el reconocimiento de las emociones faciales

El modelo de inteligencia artificial basado en redes neuronales convolucionales (CNN) fue entrenado y optimizado utilizando datasets públicos y técnicas avanzadas de Deep Learning. A continuación, se presentan los resultados obtenidos en la mejor configuración del entrenamiento:

El modelo final seleccionado se fundamenta en una arquitectura de Red Neuronal Convolucional (CNN) con múltiples capas, optimizada con el algoritmo Adam y la función de pérdida categorical_crossentropy. Se realizaron 240 épocas de entrenamiento, utilizando imágenes de 48x48 píxeles en escala de grises.

El modelo fue entrenado para clasificar siete emociones: felicidad, tristeza, ira, miedo, sorpresa, disgusto y neutralidad. Para ello, se emplearon los conjuntos de datos públicos FER2013, CK+, KDEF y RAF-DB, con un total de 299,755 imágenes. La implementación se realizó con TensorFlow y Keras en Google Colab.

Para evaluar el desempeño del modelo, se realizaron tres experimentos con diferentes configuraciones de entrenamiento tal como se muestra en la Tabla 3. El mejor modelo fue el correspondiente a Train 3, logrando un 94.62% de precisión y un F1-score promedio de 94.58%, consolidándose como la mejor configuración obtenida.

Los resultados evidencian que el modelo presentó mayor precisión en la detección de felicidad (97.50%) y tristeza (95.38%), mientras que la clasificación de neutralidad (91.33%) mostró un margen de mejora, debido a su similitud con otras emociones.

A. Comprobación de la Eficacia del Modelo de Inteligencia Artificial en el Reconocimiento Emocional en Contextos Educativos de Perú y Ecuador

La Figura 7 se proyectó en forma general la arquitectura operativa del sistema, desarrollada en Python mediante la integración de Captura de Pantalla en Tiempo Real: Utilizando PyWin32, el sistema captura las interacciones visuales, la librería OpenCV y un modelo preentrenado basado en redes neuronales profundas. La estructura funcional se descompone en las siguientes fases:

Algoritmo 1: Detecta los rostros en tiempo real utilizando la arquitectura de red neuronal configurada con archivos prototxt y caffeModel.

Algoritmo 2: Clasifica las emociones en categorías predefinidas mediante un modelo especializado (modelEFEC6.h5).

Algoritmo 3: Procesa, organiza y almacena los datos emocionales para su análisis posterior.

Análisis y Optimización del Desempeño: Los datos recolectados son procesados en tiempo real, permitiendo ajustar los algoritmos para maximizar la precisión y relevancia en el contexto educativo.

B. Análisis Comparativo de Reconocimiento Emocional entre Ecuador y Perú

Las figuras 9 y 10 presentan el análisis del desempeño del sistema de reconocimiento facial en reuniones remotas realizadas en Ecuador y Perú, evaluando tanto la interfaz del software como la distribución de emociones básicas clasificadas en positivas y negativas.

Figura 9: Reconocimiento Emocional en Ecuador

En la reunión realizada en Ecuador, participaron aproximadamente 14 estudiantes. Sin embargo, de este grupo, solo alrededor de 8 encendieron sus cámaras web, lo que limitó el análisis emocional únicamente a los participantes visibles en la interfaz del software (Figura 9a). El software identificó emociones individuales con alta precisión, destacando etiquetas como Neutral (98%) y Happy (64%).

En el gráfico de barras (Figura 9b), se excluyó la emoción Neutral debido a su alta frecuencia en comparación con otras emociones. Incluirla habría generado un sesgo significativo en la visualización estadística, dificultando un análisis equilibrado. La distribución muestra una marcada predominancia de emociones negativas, siendo Anger (1577) la más frecuente, seguida por Sadness (402). Las emociones positivas, como Happy (132) y Surprise (194), aparecen en menor proporción, lo que podría reflejar factores ambientales, culturales o situacionales específicos del contexto.

Figura 10: Reconocimiento Emocional en Perú

En la reunión realizada en Perú, participaron aproximadamente 8 estudiantes, de los cuales un máximo de 5 encendió sus cámaras web. Esta limitación en el número de participantes visibles impactó la cantidad total de datos procesados por el sistema (Figura 10a). Al igual que en Ecuador, el gráfico de barras (Figura 10b) excluyó la emoción Neutral para garantizar un balance adecuado en la representación de las emociones restantes.

La distribución emocional en Perú mostró un mayor balance entre emociones positivas y negativas en comparación con Ecuador. La emoción Happy (765) fue la más predominante, lo que sugiere un entorno emocional más positivo. Sin embargo, emociones negativas como Sadness (765) y Anger (568) también fueron significativas, mientras que Disgust (82) y Fear (94) aparecieron en menor frecuencia.

Perú presenta una distribución emocional mas equilibrada, con un 30.14% de emociones positivas como Happy, frente al 5.61% en Ecuador. En contraste, Ecuador muestra una fuerte prevalencia de emociones negativas, destacando Anger con un 67.08%, muy por encima del 22.38% muy por encima del 22.38% registrado en Perú. En resumen, Perú muestra un panorama emocional más equilibrado y positivo, mientras que Ecuador refleja una fuerte prevalencia de emociones negativas, especialmente la ira (Anger).

IV. DISCUSIÓN

En esta sección se comparan estudios recientes sobre el reconocimiento de emociones faciales mediante redes neuronales convolucionales (CNN), destacando que el modelo propuesto supera notablemente a los modelos revisados en términos de precisión, consistencia y generalización.

Éste alcanzó una precisión del 94.62% y un F1-score promedio de 94.58%, posicionándose como uno de los más

efectivos entre los evaluados, superando al modelo [16], que logró un 0.989 en el dataset CK+; sin embargo, su precisión disminuyó significativamente a 0.741 en FER2013, mostrando una capacidad limitada de generalización. Por otro lado, [23] combinaron múltiples datasets y obtuvieron una precisión máxima de 0.9232, ligeramente inferior a nuestro modelo. Asimismo, [19] implementaron un enfoque basado en explicabilidad en IA (XAI), alcanzando 0.92 en un dataset propio, pero sin superar el rendimiento general de nuestro enfoque. [5] relaciona la IA con la capacidad humana para clasificar gestos emocionales. Nuestro modelo valida esta teoría al utilizar redes neuronales para mapear gestos faciales en emociones específicas, obteniendo alta precisión en felicidad (97.50%) y tristeza (95.38%).

Otros estudios presentan limitaciones similares. Por ejemplo, [24] combinaron AffectNet y RAF-DB con un modelo preentrenado, logrando una precisión de 0.8337, notablemente inferior, mientras que [6], utilizando MobileNet y PoseNet, obtuvieron 0.844, clasificando solo seis emociones; aunque ambos enfoques son competitivos, no logran igualar la amplitud ni la precisión de nuestro modelo.

El uso de combinaciones de datasets como FER2013, CK+, KDEF y RAF-DB permitió al modelo propuesto, alcanzar una robusta capacidad de generalización, superando a estudios como [25], que se enfocaron en tareas específicas, como la clasificación de felicidad, con precisiones de 0.930 y 0.960. Ello evidencia que el enfoque propuesto no solo es más versátil, sino también más eficaz al clasificar un rango amplio de emociones.

En conclusión, el modelo propuesto se destaca por su alto desempeño que alcanzó un accuracy del 94.62%, evidenciando una alta eficacia en la clasificación de emociones faciales, superando a otros enfoques recientes al combinar precisión, consistencia y generalización, lo que lo convierte en una solución robusta y confiable en el reconocimiento de emociones faciales.
Otros autores [4], destacan la importancia de Deep Learning en la clasificación de emociones y su relación con la inteligencia artificial. Según estos autores, la capacidad de los modelos de IA para mejorar la precisión en el reconocimiento emocional depende de un entrenamiento continuo con datos diversos. En nuestro caso, se ha demostrado que el modelo entrenado con un conjunto más amplio de datos (Train 3 con 299,755 imágenes) alcanzó una precisión del 94.62%, lo que confirma la hipótesis de estos investigadores sobre la importancia del volumen de datos en la mejora del rendimiento de la IA.

Los autores [1, 26], argumentan que las expresiones faciales representan mecanismos que con ayuda de los algoritmos de visión artificial transmiten emociones facilitando las sesiones de aprendizaje educativo. Nuestro estudio respalda esta afirmación, ya que el modelo fue capaz de detectar variaciones en la expresión emocional de los estudiantes según el contexto educativo y el país.

Según [12], identifican que la ansiedad y la depresión pueden afectar la adherencia y el rendimiento académico en programas de educación en línea. Los hallazgos en Ecuador sugieren una posible correlación entre emociones negativas y el ambiente de aprendizaje; reforzando la necesidad de intervenciones pedagógicas basadas en reconocimiento emocional.

Los autores [8], sugieren que las emociones juegan un papel crucial en la adquisición del conocimiento y que los docentes deben emplear metodologías didácticas innovadoras para mejorar la interacción en clases virtuales. El sistema desarrollado, ofrece una herramienta que permite a los docentes monitorear el estado emocional de los estudiantes y adaptar su enseñanza en función de los resultados emocionales detectados.

A su vez es aplicable en entornos educativos, facilitando la evaluación del estado emocional de los estudiantes en clases remotas, brindando un potencial apoyo en el desarrollo de estrategias pedagógicas adaptativas. Por ello a partir de la comparación con investigaciones previas, podemos afirmar que:

La arquitectura propuesta basada en CNN es altamente efectiva en la clasificación de emociones faciales, con una precisión superior a estudios previos en la literatura [23, 24].

El impacto de las emociones en la educación es un factor clave, como lo sugieren [12, 1],de esta manera, el estudio aporta evidencia empírica al analizar diferencias emocionales en el entorno virtual entre estudiantes de Ecuador y Perú; en Perú destacan las emociones positivas que podrían relacionarse con el enfoque intercultural de su sistema pedagógico basado en el respeto y la empatía, valores que estarían generando bienestar subjetivo en el estudiante lo que se trasluce en emociones positivas; mientras que, en Ecuador a pesar de priorizar la educación intercultural bilingüe como derecho fundamental quizá se está dejando de lado ese toque ético y sensible lo cual se refleja en emociones negativas [30]. Esto lleva a la reflexión de que la experiencia emocional que genera el aprendizaje es compleja y requiere que los docentes propicien ambientes de aprendizaje sensibles; sin dejar de lado la interculturalidad; a través de la ecología de saberes fortaleciendo ambientes de aprendizaje virtual efectivos [28, 29].

Asimismo, la incorporación de múltiples datasets mejora la precisión del modelo, validando las observaciones [6] [24] sobre la necesidad de un entrenamiento con datos variados para aumentar la generalización del sistema. Estos hallazgos refuerzan la importancia del reconocimiento de emociones en entornos educativos y sugieren que la IA puede ser una herramienta clave para optimizar la enseñanza en clases virtuales. Como futuras líneas de investigación, se propone explorar técnicas de Transfer Learning con modelos como ResNet y Vision Transformers, así como la integración de Inteligencia Artificial Explicable (XAI) para mejorar la interpretación de los resultados emocionales.

El estudio tuvo como limitación la dificultad para identificar de manera única cada rostro, lo que impidió asociar las emociones de forma personalizada, limitando la posibilidad de realizar un seguimiento individualizado y profundo del estado emocional. Aunque el algoritmo 1 optimiza la calidad de imágenes en condiciones de iluminación extrema y sobreexposición altamente oscura; su aplicación introduce una latencia que puede afectar el desempeño en tiempo real, especialmente en dispositivos de cámara web de baja resolución dependiendo del ambiente.

REFERENCIAS

[1] Benitez-Quiroz, C. F., Srinivasan, R., & Martinez, A. M. (2018). Facial color is an efficient mechanism to visually transmit emotion. Proceedings of the National Academy of Sciences of the United States of America, 115(14), 3581–3586. https://doi.org/10.1073/PNAS.1716084115

[2] Darwin, C. (1899). "The Expression of Emotion in Man and Animals," Gutenberg. Produced by Charles Keller and David Widger. Available: https://www.gutenberg.org/files/1227/1227-h/1227-h.htm

[3] Goleman, B. (2021). Psicología Oscura 6 libros en 1: Introducción a la Psicología, Como analizar a las Personas, Manipulación, Persuasión, Secretos de la Psicología Oscura, Inteligencia Emocional y TCC, Abuso Emocional y Narcisista. Publicado: Amazon. https://psicologiaymente.com/biografias/daniel-goleman

[4] Mueller, J. P., & Massaron, L. (2019). Deep Learning For Dummies. Hoboken, NJ: John Wiley & Sons, Inc.

[5] Gokani, J. (2017, August 4). The Evolution of Banking: AI | MS&E 238 Blog. Stanford University MS&E 238 Blog. https://mse238blog.stanford.edu/2017/08/jgokani/the-evolution-of-banking-ai/

[6] Artanto, H., & Arifin, F. (2023). Emotions and gesture recognition using affective computing assessment with deep learning. IAES International Journal of Artificial Intelligence, 12(3), 1419. https://doi.org/10.11591/ijai.v12.i3.pp1419-1427

[7] Alsemawi, M. R. M., Mutar, M. H., Ahmed, E. H., Hanoosh, H. O., & Abbas, A. H. (2023). Emotions recognition from human facial images based on fast learning network. Indonesian Journal of Electrical Engineering and Computer Science, 30(3), 1478-1487. http://doi.org/10.11591/ijeecs.v30.i3.pp1478-1487

[8] Bardón, A. Á., & Andrade, A. G. (2021). El papel de las emociones en el aula virtual. The Conversation. https://theconversation.com/el-papel-de-las-emociones-en-el-aula-virtual-168435?utm_source=clipboard&utm_medium=bylinecopy_url_button

[9] Castro, Á. B. C. (2021). Motivación como estrategia para el desarrollo emocional en el proceso de aprendizaje. Dialnet. https://dialnet.unirioja.es/servlet/articulo?codigo=8217200

[10] Beltrán Baquerizo, G. E., Amaiquema Márquez, F. A., & López Tobar, F. R. (2020). La motivación en la enseñanza en línea. Revista Conrado, 16(75), 316-321.

[11] Gavotto Nogales, O. I., Castellanos Pierra, L. I.: "Las emociones negativas vividas por los maestros en las clases virtuales en tiempos de pandemia." Revista Iberoamericana para la Investigación y el Desarrollo Educativo, vol. 12, no. 23, pp. e260 (2021). Available: https://doi.org/10.23913/ride.v12i23.1006

[12] Ma L-C, Liu J, Jiao C-L, Du S-Y, Zhang R-L, Ding X-J, Shi X-Y, Wang Y. Exploring the effects of an online learning platform in stage III cardiac rehabilitation for individuals with coronary heart disease: Randomized controlled study. Medicine Baltimore. (2024);103(37). doi:10.1097/MD.0000000000039497

[13] Bachtsis, R., Perifanou, M., & Economides, A. A. (2024). Challenges Faced by Students with Special Needs in Primary Education during Online Teaching. Education Sciences, 14(3), 220. https://doi.org/10.3390/educsci14030220

[14] Zarkadoulas, D., & Virvou, M. (2024). Emotional expression in mathematics e-learning using emojis: A gender-based analysis. Intelligent Decision Technologies, 18(2), 1181–1201. https://doi.org/10.3233/idt-240170

[15] Lee, C., Manfreda, A., Bisson, A., Scerri, M., & Presbury, R. (2023). "Band-aid teaching": Exploring emergency remote teaching experiences of hospitality educators through an affective lens. Journal of Hospitality and Tourism Management, 57, 327–338. https://doi.org/10.1016/j.jhtm.2023.11.001

[16] Hao, M., Yuan, F., Li, J., & Sun, Y. (2023). Facial expression recognition based on regional adaptive correlation. IET Computer Vision, 17(4), 445–460. https://doi.org/10.1049/cvi2.12179

[17] Zhao, X., & Wang, D. (2023). Domain-Specific L2 Grit, Anxiety, Boredom, and enjoyment in online Chinese learning. The Asia-Pacific Education Researcher. https://doi.org/10.1007/s40299-023-00777-3

[18] Zatarain Cabada, R., Barrón Estrada, M. L., & Cárdenas López, H. M. (2019). Reconocimiento multimodal de emociones orientadas al aprendizaje. Research in Computing Science, 148(7), 153-165.

[19] Jiang, X., Harun, S. N., & Liu, L. (2023). Explainable artificial intelligence for ancient architecture and lacquer art. Buildings, 13(5), 1213. https://doi.org/10.3390/buildings13051213

[20] Barrionuevo, C., Ierache, J. S., & Sattolo, I. I. (2020). Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística. https://sedici.unlp.edu.ar/handle/10915/114089

[21] D. Lundqvist, A. Flykt, and A. Ohman, "The Karolinska directed emotional faces (KDEF)," CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, pp. 91–630, 1998, Accessed: Jul. 08, 2024. [Online]. Available: https://kdef.se/faq/using-and-publishing-kdef-and-akdef

[22] J. A. Tiznado Ubillús et al., "Algorithms used for facial emotion recognition: a systematic review of the literature," EAI Endorsed Trans Pervasive Health Technol, vol. 9, no. 1, p. 2023, Oct. 2023, doi: 10.4108/EETPHT.9.4214.

[23] Huang, T.-R., Hsu, S.-M., & Fu, L.-C. (2023). Data augmentation via face morphing for recognizing intensities of facial emotions. IEEE Transactions on Affective Computing, 14(2), 1228–1235. https://doi.org/10.1109/TAFFC.2021.3096922

[24] Ozdamli, F., Aljarrah, A., Karagozlu, D., & Ababneh, M. (2022). Facial recognition system to detect student emotions and cheating in distance learning. Sustainability, 14(20), 13230. https://doi.org/10.3390/su142013230

[25] Gupta, S., Kumar, P., & Tekchandani, R. K. (2022). Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. Multimedia Tools and Applications, 82(8), 11365–11394. https://doi.org/10.1007/s11042-022-13558-9

[26] Huang, Z., Chiang, C., Chen, J., Chen, Y., Chung, H., Cai, Y., & Hsu, H. (2023). A study on computer vision for facial emotion recognition. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-35446-4

[27] Aznarte, J. L., Pardos, M. M., & Lacruz López, J. M. (2022). On the use of facial recognition technologies in university: The uned case. [Sobre el uso de tecnologías de reconocimiento facial en la universidad: El caso de la uned] RIED-Revista Iberoamericana De Educacion a Distancia, 25(1), 261-277. doi:10.5944/ried.25.1.31533

[28] Marconi, A. C. (2013). Educación e interculturalidad en el Perú contemporáneo. Educación, 19, 40–45. https://doi.org/10.33539/educacion.2013.n19.1018

[29] Freire, E. E. E., & Leyva, N. V. L. (2020). Educación intercultural en el Ecuador: Una revisión sistemática. https://www.redalyc.org/journal/280/28064146018/html/

[30] Fuentes-Vilugron, G., Ortega, K. E. A., Bizama-Colihuinca, K., & Zapata, V. Z. (2023). Prácticas Pedagógicas y Educación Intercultural. Desafíos epistemológicos para la escolaridad en contextos. ResearchGate. https://www.researchgate.net/publication/375484604_Practicas_Pedagogicas_y_Educacion_Intercultural_Desafios_epistemologicos_para_la_escolaridad_en_contextos_indigenas

APÉNDICE:

Se adjunta el documento de autorización para la captura y tratamiento de datos personales, voz, Imagen y biométricos faciales:
Documento autorización:
https://drive.google.com/file/d/16_IGSAAM5TTsbEQaSk-IGzOvE9e1GpCF/view?usp=sharing

Google Formulario:
https://drive.google.com/file/d/13ejKrkAUvqly25eaZ8bqkDljlV1U0Y_y/view?usp=sharing

Dataset Emociones Faciales: Las dataset de las emociones han sido publicado que no incluye imágenes solo datos, para futuras investigaciones: https://www.kaggle.com/datasets/josetiznadoubillus/emotion-face-per-ecuador