

Prototipo de Detección de Emociones con Inteligencia Artificial

Emotion Detection Prototype with Artificial Intelligence

J.D. Morocho Fernández¹[ID 0009-0007-9809-3426](#), D. R. Guzmán,²[ID 0009-0006-2522-8898](#), L.A. Morocho Sumba³[ID 0009-0001-6662-5703](#), L.M.Calle Crespo⁴[ID 0000-0001-9643-6789](#)
david.guzman.est@tecazuay.edu.ec, jonnathan.morocho.est@tecazuay.edu.ec, luis.morocho.est@tecazuay.edu.ec,
lorena.calle@tecazuay.edu.ec

¹²³⁴Instituto Tecnológico Superior del Azuay, Cuenca/Azuay, Ecuador

DOI 10.36500/atenas.4.006

Resumen

Resumen: Este paper describe el desarrollo de un prototipo inteligente diseñado para detectar emociones en tiempo real mediante técnicas de Deep Learning, con el objetivo de contribuir al bienestar emocional de los usuarios. El sistema analiza expresiones faciales captadas por una cámara para identificar emociones como felicidad, tristeza, ira, miedo, sorpresa y asco. Los resultados se almacenan en una base de datos, junto con actividades realizadas por el usuario, lo que permite estudiar la relación entre emociones y contexto diario (hora, día y actividad). El prototipo utiliza herramientas como TensorFlow, Keras, OpenCV y DeepFace para el procesamiento y análisis de imágenes; además, de una base de datos MySQL y un dashboard interactivo en Power BI para la visualización de resultados. Entre sus aportes destaca la recomendación automática de música personalizada según la emoción detectada, lo que refuerza el vínculo entre estado emocional y música. En este contexto, el modelo fue entrenado con bases de datos públicas y registros propios, a través de la aplicación de técnicas de balanceo de clases (SMOTE), normalización de datos y redes neuronales con IA. Se obtuvo una precisión promedio del 40%, con mejores resultados en emociones como “sorpresa” y “felicidad”. La validación realizada por expertos en psicología y tecnología evidenció un 75% de efectividad en las recomendaciones musicales propuestas por el sistema.

Abstract

Abstract: This paper presents the development of an intelligent prototype designed to detect emotions in real time using deep learning techniques, with the aim of contributing to users’ emotional well-being. The system analyzes facial expressions captured by a camera to identify emotions such as happiness, sadness, anger, fear, surprise, and disgust. The results are stored in a database alongside user activities, enabling the study of relationships between emotions and daily context (time, day, and activity). The prototype employs tools such as TensorFlow, Keras, OpenCV, and DeepFace for image processing and analysis, as well as a MySQL database and an interactive Power BI dashboard for results visualization. A notable feature is the automatic recommendation of personalized music based on the detected emotion, reinforcing the connection between emotional state and music. The model was trained using public datasets and proprietary records, applying class balancing techniques (SMOTE), data normalization, and artificial neural networks. An average accuracy of 40% was achieved, with higher performance for emotions such as “surprise” and “happiness.” Validation by experts in psychology and technology indicated a 75% effectiveness in the system’s music recommendations.

Palabras Claves – Emociones, Psicología, Biometría, Inteligencia Artificial, Aprendizaje Profundo, Base de Datos
Keywords – Emotions, Psychology, Biometry, Artificial Intelligence, Deep Learning, Databases

I. INTRODUCCIÓN

El reconocimiento de emociones ha cobrado relevancia en los últimos años debido a su impacto en múltiples ámbitos, como la toma de decisiones, la interacción social y el bienestar emocional. En la era digital, la integración de la inteligencia artificial en este campo ofrece nuevas oportunidades para el análisis y comprensión de los estados emocionales, con aplicaciones en áreas como la psicología, la educación y el entretenimiento. En este contexto, un prototipo capaz de detectar emociones y recomendar música personalizada podría mejorar la experiencia del usuario al adaptar la música a su estado de ánimo.

La relación entre música y emociones ha sido estudiada desde la antigüedad. Hace más de 2500 años, Pitágoras utilizaba escalas y acordes específicos para inducir el equilibrio mental, recomendando melodías para armonizar estados emocionales adversos. Aristóteles observó que ciertos ritmos y melodías fortalecen cuerpo y espíritu; mientras que, Platón comparaba la música con la gimnasia para el alma. En la modernidad, Descartes sostuvo que la música tenía como propósito complacer y provocar emociones (Schweppe & Schweppe, 2010). Estas reflexiones evidencian cómo la música ha sido utilizada históricamente como un mecanismo para modular el estado anímico de las personas.

El presente trabajo propone el desarrollo de un detector de emociones basado en técnicas de Deep Learning, capaz de analizar expresiones faciales en tiempo real para identificar emociones predominantes y su distribución. Este prototipo incorpora una base de datos para almacenar los resultados obtenidos y relacionarlos con actividades específicas reportadas por los usuarios. Su implementación no solo permite estudiar el impacto emocional en la vida diaria; sino que, también abre la posibilidad de integrar aplicaciones innovadoras, como sistemas de recomendación musical personalizados.

Los sistemas de recomendación son herramientas que generan sugerencias basadas en las preferencias y experiencias de los usuarios. En este sentido, la combinación de reconocimiento facial con algoritmos de recomendación musical podría proporcionar una experiencia más personalizada y significativa.

El objetivo principal de este proyecto fue desarrollar una herramienta capaz de identificar patrones emocionales en tiempo real y explorar sus implicaciones en actividades cotidianas, lo que contribuye a un bienestar emocional más profundo. Como objetivo secundario, se busca analizar cómo la música puede influir en la regulación emocional al adaptarse a los estados de ánimo detectados.

La hipótesis de este estudio plantea que es posible identificar emociones a partir del análisis de expresiones faciales y que dicha información, complementada con datos sobre las actividades personales de los individuos, puede ser aplicada al desarrollo de sistemas orientados a promover el bienestar emocional. Para ello, se emplean algoritmos de aprendizaje profundo en el procesamiento de imágenes faciales, junto con una base de datos estructurada que permite almacenar y analizar la información recopilada de manera eficiente.

Bajo esta hipótesis surge la pregunta de investigación ¿En qué medida la identificación automática de emociones mediante el análisis de expresiones faciales, apoyada en el contexto de las actividades personales de los usuarios, puede contribuir al diseño de sistemas que fomenten el bienestar emocional?

En este contexto, se consideran como variables independientes las expresiones faciales detectadas a través de algoritmos de aprendizaje profundo y los datos relacionados con las actividades personales de los usuarios. La variable con la que se pretende responder la pregunta es la emoción identificada.

Este enfoque no solo facilita la identificación de emociones; sino que, también establece una conexión directa con las actividades específicas de los usuarios, lo que representa una innovación en el campo del reconocimiento emocional. Finalmente, este artículo analiza los posibles alcances y aplicaciones sociales de esta tecnología.

II. MARCO TEÓRICO

En la fase inicial del proyecto se llevó a cabo una revisión teórica para poder identificar las emociones; según Ekman (2003) quien propone las siguientes emociones: felicidad, tristeza, ira, miedo, sorpresa y asco, argumenta que son universales y se expresan de forma similar a través de expresiones faciales en todas las culturas humanas, la detección de emociones y la influencia que la música puede ejercer sobre el estado anímico de los individuos. Diversos estudios respaldan que la música desempeña un papel esencial en la vida cotidiana, no solo como forma de entretenimiento, sino como medio expresivo que permite manifestar y modificar emociones.

De acuerdo con Fernández (2019), la música posee un valor significativo tanto a nivel individual como social. No solo transmite emociones; sino que, también comunica significados profundos, representa valores culturales y tiene la capacidad de recrear experiencias emocionales. En palabras del autor, la música forma parte intrínseca de la existencia humana y puede concebirse como una entidad dinámica con la capacidad de influir en la vida emocional de las personas (p. 27).

Este potencial emocional y comunicativo de la música ha sido aprovechado por la musicoterapia, una disciplina terapéutica que utiliza la música como herramienta para facilitar el desarrollo personal. Según la Federación Mundial de Musicoterapia (1996), esta práctica busca fortalecer habilidades latentes o recuperar funciones deterioradas en los individuos, a través del fomento de su integración, tanto a nivel emocional como social.

Tecnología y herramientas

Durante la etapa de desarrollo del prototipo, se identificaron y seleccionaron diversas herramientas y bibliotecas mediante búsquedas en línea y revisión de literatura técnica. A continuación, se describen las principales tecnologías utilizadas.

TensorFlow: es una biblioteca de código abierto desarrollada por Google que proporciona diversas herramientas para la creación y entrenamiento de modelos de aprendizaje automático a gran escala (TensorFlow, 2023). Por su parte, **Keras** es una API de alto nivel que funciona sobre TensorFlow, utilizada en este prototipo para facilitar la construcción y entrenamiento de redes neuronales de forma más sencilla y rápida (Keras, 2023).

NumPy (Numerical Python): es una biblioteca esencial para la computación científica en Python, que ofrece estructuras de datos eficientes, principalmente arreglos multidimensionales, junto con una variedad de herramientas matemáticas especializadas, incluyendo funciones de álgebra lineal, que facilitan el procesamiento y análisis numérico avanzado (Harris et al., 2020).

PIL: es una biblioteca que ofrece soporte para una gran variedad de formatos de imagen, lo que permite abrir, modificar y guardar archivos de forma sencilla. Esta herramienta resulta fundamental para el procesamiento de imágenes en tareas relacionadas con visión computacional y análisis visual (Pillow, 2024).

Matplotlib: es una biblioteca de Python destinada a la generación de gráficos bidimensionales, que permite representar y analizar visualmente conjuntos de datos de forma clara y precisa. Gracias a su estructura flexible y altamente configurable, se ha convertido en una herramienta fundamental en tareas de investigación, exploración de datos y desarrollo de soluciones científicas (Matplotlib, 2024).

Mysql.connector: es un conector oficial de Python para conectarse a bases de datos MySQL. Permite ejecutar sentencias SQL desde un script Python.

Deepface: es una biblioteca de código abierto desarrollada en Python para el reconocimiento facial y el análisis de atributos faciales como edad, género, emociones y etnia. Integra múltiples modelos preentrenados de aprendizaje profundo, lo que permite realizar análisis faciales de forma eficiente y con mínima configuración, lo que facilita su incorporación en aplicaciones de visión computacional (Serengil & Ozpinar, 2020; Serengil, 2024).

Streamlit: es una biblioteca de Python diseñada para simplificar el desarrollo de aplicaciones web interactivas; especialmente, orientadas a la visualización de datos y la implementación de modelos de aprendizaje automático en entornos de ciencia de datos (Streamlit, 2024).

OpenCV: es una biblioteca de visión por computadora que permite el procesamiento de imágenes y vídeos, siendo ampliamente utilizada en tareas como detección de objetos (OpenCV, 2024). En el presente proyecto, resulta fundamental para la identificación y manipulación de rostros durante el análisis emocional.

III. METODOLOGÍA Y CÁLCULOS

La metodología utilizada en el desarrollo del presente proyecto se basa en un enfoque cuantitativo, aplicado y experimental, orientado a la construcción de un prototipo funcional de detección de emociones en tiempo real. Con el fin de responder a las preguntas de investigación planteadas, se diseñó una metodología estructurada en cinco fases principales. Cada una de estas fases está alineada con una o más preguntas de investigación (Research Questions, RQ), lo que permite abordar de forma sistemática el desarrollo e implementación del prototipo de detección de emociones y su vínculo con la música. La estructura propuesta se muestra en la Figura 1 y se detalla a continuación:

Figura 1

Arquitectura del Proyecto



Nota. La figura muestra la arquitectura general del sistema desarrollado, dividido en cinco etapas.

Fase 1: se aplicó la recopilación de datos, lo cual se relaciona con la pregunta de investigación RQ1, ¿Cómo se diseñó el dataset para el prototipo de detección de emociones?

El primer paso en la arquitectura fue la recolección de un conjunto de datos desde la plataforma web Kaglee, donde se descargó un dataset en el que las imágenes estaban divididas según sus emociones. Una vez recopiladas, las imágenes fueron clasificadas según la emoción más prominente,

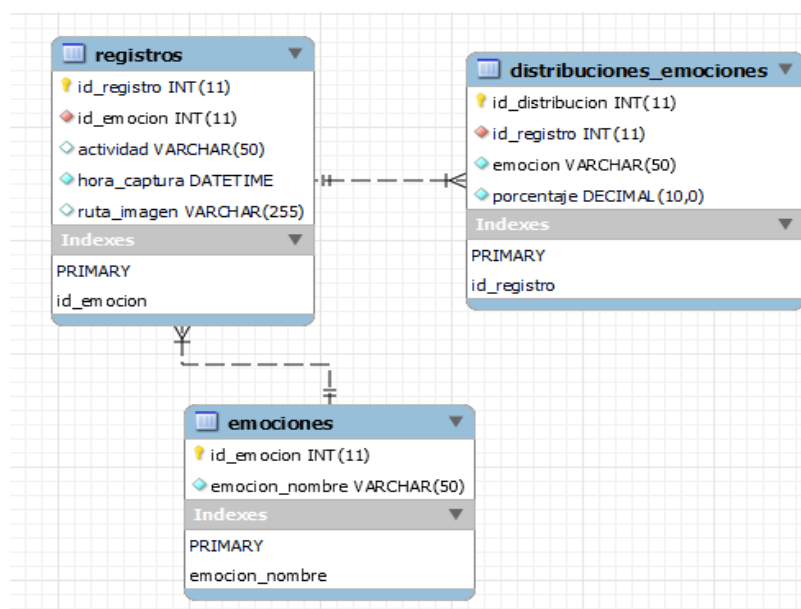
seguida de las demás emociones presentes en cada imagen. Este proceso facilitó la creación de la base de datos y optimizó el etiquetado para su posterior utilización.

Fase 2: se procede con la creación de la base de datos utilizando MySQL. Esta es una base de datos relacional, la cual fue elegida por su capacidad de integrarse eficientemente con la aplicación de detección de emociones. Se empleó el lenguaje SQL para estructurar, asegurando que cumpliera con los principios de normalización para optimizar su diseño y evitar redundancias.

La base de datos se llamó “detector emociones” que engloba tres tablas principales que se muestra a continuación:

Figura 2

Modelo Relacional



Nota. Se muestra la estructura de la base de datos utilizada en el prototipo de detección de emociones, compuesta por tres tablas principales.

A. Tabla de emociones (emociones)

Esta tabla almacena las emociones disponibles, identificadas por un ID único. Cada emoción se registra en el campo **emocion_nombre**, el mismo que sirve para evitar duplicados.

B. Tabla de registros (registros)

Esta tabla guarda información sobre las imágenes procesadas, asociando cada registro a una emoción específica mediante una clave foránea (`id_emocion`). También incluye información adicional como la actividad registrada, la hora de captura (por defecto, la hora actual) y la ruta donde se encuentra la imagen.

C. Tabla de distribuciones emocionales (distribuciones_emociones)

Esta tabla almacena las probabilidades de las emociones detectadas en cada imagen. Cada registro está vinculado a la tabla **registros** mediante una clave foránea (`id_registro`). También incluye el porcentaje correspondiente a cada emoción.

Fase 3: se diseñó la aplicación “Detección de emociones”, la cual va relacionada con la pregunta de investigación RQ2, ¿Qué técnicas se utilizan para el reconocimiento de emociones?

En esta fase, se desarrolló la aplicación en Visual Studio Code utilizando el lenguaje de programación Python, con el objetivo de detectar las emociones faciales en tiempo real de los usuarios. La solución combina técnicas de procesamiento de imágenes con algoritmos de inteligencia artificial basados en aprendizaje profundo. El flujo del proceso se detalla a continuación:

Captura de rostro

La función `procesar_emociones` constituye la función principal del prototipo de análisis emocional de imágenes en el aplicativo desarrollado. Su propósito es detectar rostros humanos en un fotograma de imagen proporcionado y, a partir de estos, realizar un análisis de emociones utilizando el modelo DeepFace.

Esto se logra mediante un modelo pre entrenado, basado en un clasificador Haar Cascade, el cual convierte el fotograma de formato BGR a escala de grises, ya que este formato es requerido para el análisis emocional.

Análisis emocional

Una vez identificado el rostro, se extrae la región relevante de la imagen y se analiza con el modelo DeepFace. Este modelo, basado en redes neuronales profundas, es capaz de identificar emociones humanas a partir de características faciales. Las emociones identificables abarcan felicidad, tristeza, enojo, sorpresa, miedo y asco. DeepFace genera una distribución porcentual que indica la intensidad relativa de cada emoción.

Visualización de resultados

Los resultados del análisis se presentan al usuario mediante:

La emoción dominante, acompañada de un emoji representativo.

Un gráfico circular para ilustrar la distribución porcentual de las emociones detectadas, brindando una representación detallada del panorama emocional general.

Almacenamiento de Resultados

Con el objetivo de conservar un registro de los análisis realizados a cada usuario y su información correspondiente, se almacena en una base de datos MySQL, tanto la emoción dominante; así como, la distribución emocional, actividad; además, de la fecha y hora. Esta tarea se realiza mediante la función `guardar_en_mysql()`, lo que permite disponer, posteriormente, de los datos necesarios para un análisis poblacional.

Interacción con música

Relacionado con la pregunta de investigación RQ3, ¿Cuál es el impacto de la música en los estados de ánimo?

Se implementó una funcionalidad que sugiere una playlist personalizada en función de la emoción dominante detectada por el prototipo; la aplicación sugiere una playlist que ayuda a mejorar el estado emocional del usuario. Esta recomendación se genera automáticamente tras el análisis facial y se presenta al usuario mediante un enlace directo a una plataforma de streaming Spotify.

Entrenamiento del algoritmo

Para el entrenamiento del modelo de Deep Learning, se utilizaron los datos almacenados previamente en la base de datos relacional, diseñada específicamente para la aplicación de detección de emociones. Esta base de datos almacena registros detallados de los análisis efectuados, incluyendo datos sobre las emociones identificadas, la actividad realizada por los usuarios, la hora del registro y otros datos relevantes.

Para la preparación del conjunto de entrenamiento, se ejecutó una consulta SQL diseñada para extraer únicamente los campos relevantes para el modelo. Los datos obtenidos a través de esta consulta incluyen:

r.actividad: La actividad que la persona estaba realizando al momento del análisis.

r.hora_captura: la hora exacta en la que se capturó la imagen y se detectó la emoción.

e.emocion_nombre: la hora exacta en la que se capturó la imagen y se identificó la emoción.

Procesamiento de datos Temporales

Como parte del procesamiento previo al entrenamiento del modelo de Deep Learning, se aplicaron transformaciones a las variables temporales con el objetivo de facilitar su interpretación por parte del algoritmo.

Transformación de la hora del día

La hora en que fue capturada la emoción se representó en formato circular, para conservar la naturaleza cíclica del tiempo (por ejemplo, las 23:00 y las 00:00 están próximas entre sí). Para ello, se aplicaron funciones trigonométricas de seno y coseno de la hora normalizada, por medio de la generación de dos nuevas variables:

$$hora_sin = \sin(2\pi * hora/24) \quad (1)$$

$$hora_cos = \cos(2\pi * hora/24) \quad (2)$$

Día de la semana:

Se transformó el día de la semana a valores numéricos comprendidos entre 0 y 6, donde:

0 = Lunes, a 6 = Domingo.

Mes del año:

El mes fue convertido a su correspondiente valor numérico (1 para enero, 2 para febrero, ..., 12 para diciembre), permitiendo así identificar variaciones estacionales en los estados emocionales.

Identificación de fin de semana:

Se añadió una variable binaria denominada `fin_de_semana`, la cual toma el valor 1, si el día corresponde a sábado (5) o domingo (6); y, 0 en caso contrario. Esto permitió evaluar si las emociones varían en función de si es un día laboral o de descanso.

Procesamiento de Datos Categóricos

Para preparar las variables categóricas se utilizó la técnica de codificación de etiquetas (Label Encoding) mediante la clase `LabelEncoder` de la librería `scikit-learn`.

Se aplicó esta transformación a las siguientes variables:

actividad: describe la acción realizada por el usuario en el momento del análisis emocional.

emocion_nombre: representa la emoción dominante detectada por el modelo `DeepFace`.

Dado que estas variables estaban en formato de texto, el `LabelEncoder` permitió convertir cada categoría en un valor numérico entero.

Selección de Variables Predictoras y variable objetivo

Para el entrenamiento del modelo de Deep Learning, se definió un conjunto de variables predictoras (X) y una variable objetivo (y), que representa la emoción dominante que se desea predecir.

Las variables predictoras seleccionadas fueron:

actividad: codificada numéricamente, representa la acción realizada por el usuario.

hora_sin y hora_cos: transformación de la hora en componentes circulares para representar la naturaleza cíclica del tiempo.

dia_semana: número del día de la semana, donde lunes = 0 y domingo = 6.

mes: número del mes del año (1 al 12).

fin_de_semana: variable binaria que indica si el día corresponde a un fin de semana (1) o no (0).

La variable objetivo y, que se desea predecir, corresponde a la emoción dominante detectada en cada registro. Esta separación entre variables independientes y la variable dependiente es esencial para entrenar el modelo.

Escalado de Características

Normalizamos las variables predictoras para asegurar que todas las características tuvieran una escala comparable y evitar que algunas dominen el entrenamiento debido a sus valores numéricos. Para ello, se utilizó la técnica de escalado estándar, mediante la clase `StandardScaler`.

Balanceo de clases con SMOTE

Durante la exploración inicial de los datos, se identificó un desbalance en las clases emocionales; es decir, algunas emociones estaban representadas con muchos más registros que otras. Esta situación puede provocar que el modelo de Deep Learning aprenda un sesgo hacia la clase mayoritaria, lo que reduce su capacidad para predecir correctamente emociones menos frecuentes.

Para abordar este problema, se utilizó la técnica SMOTE (Synthetic Minority Over-sampling Technique), que genera ejemplos sintéticos para las clases minoritarias mediante interpolación de registros reales. Esto mejora el equilibrio en la distribución de clases sin duplicar datos, lo que resulta en un mejor rendimiento y generalización del modelo.

División de Entrenamiento y Prueba

Para evaluar el rendimiento del modelo de forma objetiva, se dividió el conjunto de datos balanceado en dos subconjuntos:

80% de los datos, se utilizaron para entrenar el modelo (X_{train} , y_{train}).

20%, se reservaron para evaluar su rendimiento (X_{test} , y_{test}).

Creación de la Red Neuronal Artificial ANN

El modelo de clasificación fue diseñado usando Keras, como una red neuronal secuencial compuesta por tres capas ocultas densamente conectadas. El objetivo del modelo es predecir la emoción dominante del usuario en función de variables contextuales.

La división se realizó de manera estratificada utilizando `train_test_split` de `sklearn`, lo que garantiza que las proporciones de las clases se mantuvieran en ambos subconjuntos, lo que evita sesgos durante la evaluación.

Características clave del Modelo

Capas ocultas: Tres capas densas con 128, 64 y 32 neuronas, respectivamente.

Función de activación: Se utilizó LeakyReLU, que permite el paso de un pequeño gradiente cuando la unidad no se activa.

Batch Normalization: Mejora la estabilidad y acelera el proceso de entrenamiento al normalizar la salida de cada capa.

Dropout (30%): Se aplicó después de cada capa para prevenir el sobreajuste (overfitting), asegurando mejor generalización.

Capa de salida: Utiliza activación softmax para predecir la probabilidad de cada emoción, siendo adecuada para un problema de clasificación multiclase.

Análisis personalizado, mediante un Dashboard

Como último paso en la metodología de nuestro proyecto, se desarrolló un dashboard personalizado en Power BI para visualizar y analizar los registros de emociones capturados por la aplicación. Este dashboard permite a los usuarios explorar los datos de manera interactiva y obtener información valiosa sobre las emociones detectadas.

Recopilación de Datos

Se registran las emociones detectadas, junto con marcas de tiempo y otras métricas relevantes, en una base de datos estructurada para facilitar su análisis.

Integración con Power BI

Mediante una integración directa, entre la base de datos y Power BI, se agiliza la importación de los datos esenciales para los gráficos interactivos:

Diseño del Dashboard

Se elabora un diseño intuitivo que incluye gráficos interactivos, como líneas de tiempo y gráficos de barras, para representar la evolución y frecuencia de las emociones a lo largo del tiempo.

Se implementan filtros que permiten a los usuarios segmentar los datos por períodos específicos o tipos de emociones, mejorando la experiencia de análisis.

Análisis Visual

El dashboard proporciona una representación visual clara de los datos, lo que facilita la identificación de patrones y tendencias emocionales, crucial para la toma de decisiones informadas.

Actualización Continua

Se integra un sistema de actualización automática para garantizar que el dashboard refleje siempre los datos más recientes, manteniendo su relevancia y utilidad en el tiempo.

Este enfoque final no solo mejora la comprensión de los datos recopilados; sino que, también potencia la capacidad de los usuarios para realizar análisis profundos y tomar decisiones basadas en información emocional precisa.

Valoración del prototipo

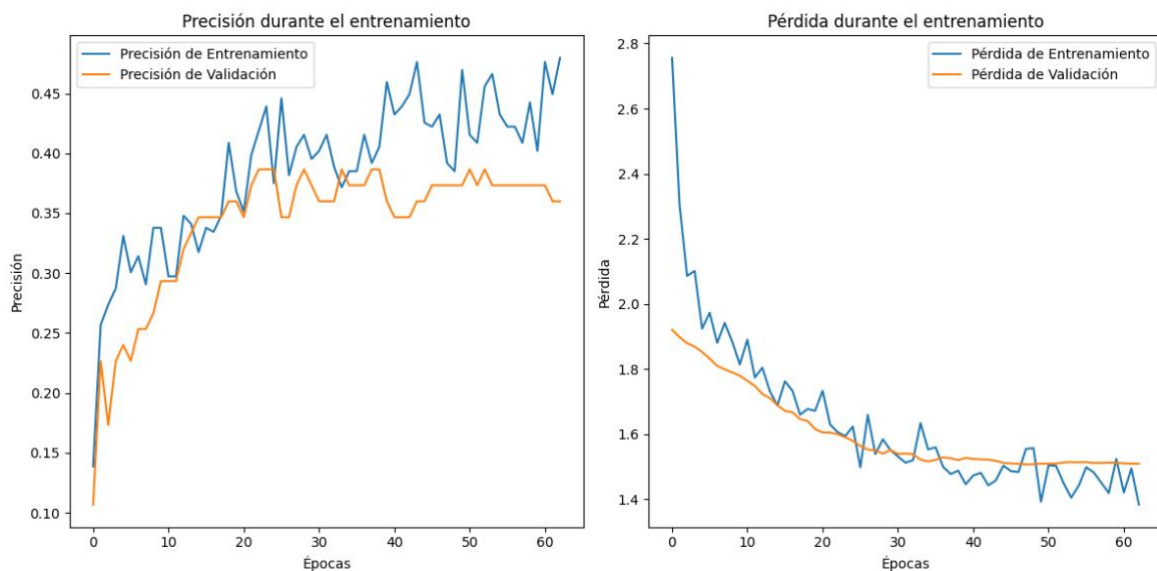
Con la finalidad de determinar la efectividad del prototipo desarrollado para la detección de emociones, se ha llevado a cabo un proceso de validación utilizando un conjunto de datos previamente etiquetado y no visto por el modelo. Esta evaluación permite analizar el desempeño del sistema mediante métricas estándar en clasificación multiclase, como la precisión por clase, la matriz de confusión y la evolución de precisión/pérdida durante el entrenamiento.

Se ilustra el comportamiento del modelo durante el proceso de entrenamiento, mostrando la evolución de la precisión y la pérdida, tanto en los datos de entrenamiento; así como, de validación. Se

evidencia una mejora progresiva hasta estabilizarse alrededor de la época 50, donde el modelo alcanza una precisión de validación de aproximadamente 0.40, consolidando así su aprendizaje, sin indicios severos de sobreajuste.

Figura 3

Evolución de la precisión y pérdida durante el entrenamiento del modelo

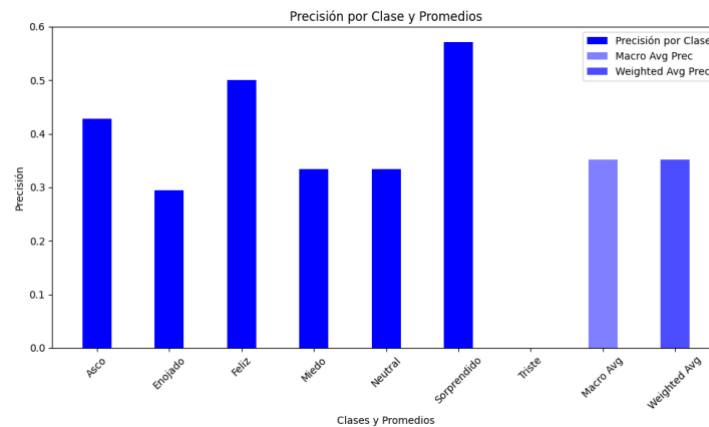


Nota. Se observa el rendimiento del modelo en entrenamiento y validación por épocas.

Como segundo lugar se muestra la precisión alcanzada por el modelo para cada una de las emociones. Se observa que la emoción "Sorprendido", obtuvo la mayor precisión (aproximadamente, 0.58), seguida de "Feliz" (0.50) y "Asco" (0.43). Por otro lado, emociones como "Enojado" y "Triste" presentaron niveles de precisión relativamente más bajos, lo cual puede estar relacionado con la similitud en patrones faciales o el desbalance en el conjunto de entrenamiento. Además, se incluye el macro promedio (0.35) y el promedio ponderado (0.35) de precisión, como referencia general del rendimiento del modelo.

Figura 4

Precisión por clase y promedios en la detección de emociones.

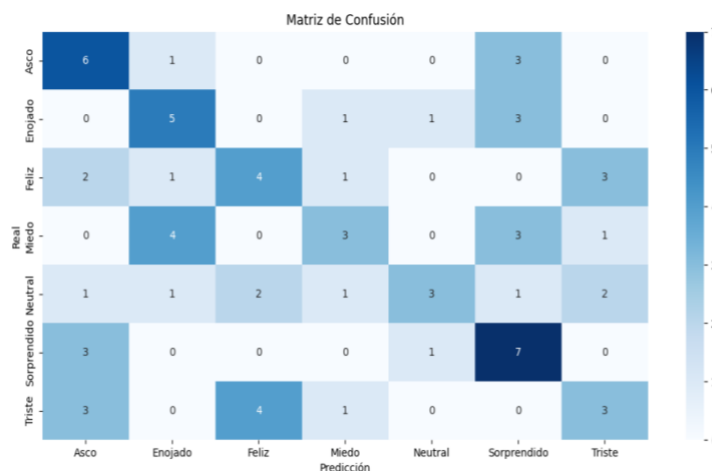


Nota. Se visualiza la precisión alcanzada para cada emoción y los promedios macro y ponderado.

A continuación, la matriz de confusión, en donde se puede observar la distribución de las predicciones realizadas por el modelo en relación con las clases reales. Este análisis permite identificar con mayor claridad las emociones que tienden a ser confundidas entre sí, como es el caso de "Triste" con "Feliz"; o, "Miedo" con "Enojado", lo cual evidencia la necesidad de seguir afinando el sistema para mejorar su capacidad discriminativa.

Figura 5

Matriz de confusión del modelo de detección de emociones



Nota. Muestra aciertos y errores en la predicción por clase, según los valores reales.

IV. RESULTADOS Y DISCUSIÓN

Se evaluó el prototipo con un grupo de cinco expertos en psicología y tecnología. Los resultados evidenciaron un 75% de efectividad en la precisión de las recomendaciones musicales basadas en emociones detectadas. Todos los participantes firmaron un consentimiento informado por escrito, en el cual autorizaron el uso de los datos recopilados con fines académicos, investigativos y de publicación. Dicho proceso cuenta con la aprobación del Instituto Superior Tecnológico Universitario del Azuay, lo que garantiza la validez y transparencia del procedimiento. Asimismo, los expertos coincidieron en que el prototipo posee un alto potencial de aplicación cotidiana, dado que permite establecer una correlación entre la actividad realizada y el estado emocional, además de confirmar que la música influye directamente en el estado de ánimo del usuario.

Las observaciones específicas de los expertos fueron:

Psicólogo experto, Cristian Quezada: resaltó la importancia de explorar métodos alternativos para detectar emociones en personas que no manifiestan expresiones faciales claras, lo que sugiere una futura línea de investigación para mejorar la sensibilidad del sistema.

Psicóloga experta, María Paz Machado: enfatizó la necesidad de explicar claramente el objetivo y fin de la actividad al usuario antes de utilizar el prototipo, para mejorar la disposición emocional y el contexto de interpretación.

V. CONCLUSIONES

En un mundo donde las emociones tienen varias aristas y poder identificarlas resulta esencial para proporcionar una experiencia personalizada y significativa, este prototipo permitió demostrar que la actividad tiene un efecto inmediato en el estado de ánimo de las personas. A futuro, esto puede contribuir para comprender mejor las emociones en relación con el entorno y, de esta manera, encontrar soluciones efectivas como la música.

La música provoca una reacción significativa en la generación de sentimientos, emociones y diversas respuestas. Por ello, desarrollar mecanismos de recomendaciones musicales basadas en emociones representa un avance en la forma en que las personas se relacionan con la música y cómo

esta puede influir positivamente en el estado de ánimo individual, lo cual constituye un aporte valioso para la salud mental.

REFERENCIAS BIBLIOGRÁFICAS

- Fernández Poncela, A. M. (2019). La autorregulación emocional de las juventudes a través de la música. *ESCENA. Revista de las Artes*, 79(1), 25–58. <https://doi.org/10.1109/LSP.2016.2603342>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. <https://www.deeplearningbook.org/>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Kacti, M. N. (2019). Music therapy in clinical psychology. En V. V. R. V. D. K. (Ed.), *Psychology and clinical therapy: Therapeutic interventions through music* (pp. 75–92). Academic Press. <https://doi.org/10.1016/B978-0-12-813574-7.00007-5>
- Kaggle. (s. f.). Facial emotion recognition dataset. Kaggle. <https://www.kaggle.com/datasets/msambare/fer2013>
- Keras. (2023, 10 de julio). Keras documentation: Introduction to Keras for engineers. Keras. Recuperado el 1 de diciembre de 2024, de https://keras.io/getting_started/intro_to_keras_for_engineers/
- Martín de Diego, I., Serrano, A., Conde, C., & Cabello, E. (2006). Técnicas de reconocimiento automático de emociones. *Teoría de la Educación: Educación y Cultura en la Sociedad de la Información*, 7(2), 92–106. <https://gredos.usal.es/handle/10366/56524>
- Matplotlib. (2024). Matplotlib documentation — Matplotlib 3.10.3 documentation. Matplotlib. Recuperado el 1 de diciembre de 2024, de <https://matplotlib.org/stable/>
- OpenCV. (2024). About OpenCV. <https://opencv.org/about/>
- Pillow. (2024). Pillow (PIL Fork) 11.2.1 documentation. Recuperado el 1 de diciembre de 2024, de <https://pillow.readthedocs.io/en/stable/>

- Rojas Espinoza, T., & Mármol Córdoba, M. F. (2024). Inteligencia artificial y su impacto en los entornos de educación superior del Ecuador. *ATENAS Revista Científica Técnica y Tecnológica*, 3(1). <https://doi.org/10.36500/atenas.3.001>
- Serengil, S., & Ozpinar, A. (2020). LightFace: A hybrid deep face recognition framework. En 2020 Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 23–27). IEEE. <https://doi.org/10.1109/ASYU50717.2020.9259802>
- Streamlit. (2024). Streamlit documentation. <https://docs.streamlit.io/>
- TensorFlow. (2023, 2 de marzo). Guide. TensorFlow. Recuperado el 1 de diciembre de 2024, de <https://www.tensorflow.org/guide>
- Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>