



**PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR
ESCUELA INFORMÁTICA E INTELIGENCIA ARTIFICIAL**

**TRABAJO DE INTEGRACIÓN CURRICULAR PREVIO A LA OBTENCIÓN DEL
TÍTULO DE INGENIERO EN TECNOLOGÍAS DE LA INFORMACIÓN.**

**PROTOTIPO DE RECONOCIMIENTO FACIAL PARA LA EVALUACIÓN
EMOCIONAL DEL PERSONAL DOCENTE DE LA ESCUELA DE INGENIERÍA,
DE LA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE IBARRA**

KENY ARIEL MAFLA PINEDA

TUTOR: FRANCKLIN IVAN RIVAS ECHEVERRIA

IBARRA – ECUADOR

FEBRERO, 2025

Ibarra, 20 de febrero de 2025

CERTIFICACIÓN TUTOR

En mi calidad de Tutor del Trabajo de titulación titulado:

PROTOTIPO DE RECONOCIMIENTO FACIAL PARA LA EVALUACIÓN EMOCIONAL DEL PERSONAL DOCENTE DE LA ESCUELA DE INGENIERÍA, DE LA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE IBARRA, presentado por el estudiante KENY ARIEL MAFLA PINEDA con cédula de ciudadanía N° 0401787650, para obtener el Título de INGENIERO EN TECNOLOGÍAS DE LA INFORMACIÓN.

Certifico que el trabajo cumple con todos los parámetros establecidos, mediante el cual el estudiante demuestra el desarrollo de competencias en el campo de conocimiento de su profesión con un nivel de argumentación coherente, para ser sometido a la evaluación por parte de los lectores.

Adicionalmente, se adjunta el certificado de porcentaje de originalidad de TURNITIN.


21/1/25, 7:40 a.m. Turnitin - Informe de Originalidad - Trabajo de titulación_Mafla Keny.docx

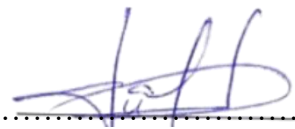
Turnitin Informe de Originalidad					
Procesado el: 21-ene.-2025 07:34 -05 Identificador: 2568249727 Número de palabras: 16145 Entregado: 1					
Trabajo de titulación_Mafla Keny.docx Por KENY ARIEL MAFLA PINEDA	<table border="1"><thead><tr><th>Índice de similitud</th><th>Similitud según fuente</th></tr></thead><tbody><tr><td>6%</td><td>Internet Sources: 4% Publicaciones: 1% Trabajos del estudiante: 5%</td></tr></tbody></table>	Índice de similitud	Similitud según fuente	6%	Internet Sources: 4% Publicaciones: 1% Trabajos del estudiante: 5%
Índice de similitud	Similitud según fuente				
6%	Internet Sources: 4% Publicaciones: 1% Trabajos del estudiante: 5%				
3% match (trabajos de los estudiantes desde 04-jul.-2024) Submitted to Pontificia Universidad Catolica del Ecuador - PUCE on 2024-07-04					
1% match (trabajos de los estudiantes desde 30-ago.-2024) Submitted to Pontificia Universidad Catolica del Ecuador - PUCE on 2024-08-30					

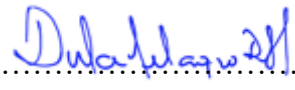
(f): _____
PhD. Francklin Rivas
TUTOR DE TRABAJO
C.C.: 175706379-5

PÁGINA DE APROBACIÓN DEL TRIBUNAL

El tribunal examinador, aprueba el presente trabajo en nombre de la Pontificia Universidad Católica del Ecuador Ibarra:

(f): 
.....
PhD. Francklin Rivas
C.C.: 175706379-5

(f): 
.....
PhD. Laura Guerra
C.C.: 1757842784

(f): 
.....
PhD. Dulce Rivero
C.C.: 1757608961

ACTA DE CESIÓN DE DERECHOS

Yo, *KENY ARIEL MAFLA PINEDA*, declaro conocer y aceptar la disposición del Art. 165 del Código Orgánico de Economía Social de los Conocimientos, Creatividad e Innovación, que manifiesta textualmente: “Se reconoce facultad de los autores y demás titulares de derechos de disponer de sus derechos o autorizar las utilidades de sus obras o prestaciones a título gratuito y oneroso, según las condiciones que determinen. Esta facultad podrá ejercerse mediante licencias libres, abiertas y otros modelos alternativos de licenciamiento o la renuncia”.

Ibarra, 20 de febrero de 2025

(f):  _____

KENY ARIEL MAFLA PINEDA

C.C.: 0401787650

AUTORIA

Yo, *KENY ARIEL MAFLA PINEDA*, portador de la cedula de ciudadanía N° 0401787650, declaro que el presente trabajo de investigación es de total responsabilidad del autor, y eximo expresamente a la Pontificia Universidad Católica del Ecuador Ibarra de posibles reclamos o acciones legales.

(f):.....

KENY ARIEL MAFLA PINEDA

C.C.: 0401787650

DEDICATORIA

A mi grandioso Dios, por darme la vida y permitirme disfrutar de esta etapa tan bonita, enriquecedora y llena de muchos momentos, tanto buenos como malos, y por estar siempre a mi lado.

A mi amada madre, Karina, quien me trajo al mundo y me ha llenado de su amor, sabiduría, esencia y paciencia infinita. Mamita, esto se lo dedico a usted, por siempre confiar en mí, por sus grandes enseñanzas, por su sacrificio y por estar siempre para mí, usted es un pilar fundamental para que hoy esté cumpliendo un sueño más.

A mi amado padre, Gary, por ser el ejemplo de un gran hombre, por inculcarme lo máspreciado, el ejemplo. Gracias por su amor, por sus grandes sacrificios y por estar siempre pendiente de mí, asegurándose de que nunca me falte nada. Papito, esta dedicación también es para usted, por ser otro pilar esencial para este sueño.

A mi querido hermano, Jhordan, por ser siempre un apoyo, por estar ahí, por escucharme y por ser ese hermano mayor que me guía, me aconseja y me da su apoyo. Ñaño, eres otro pilar de este nuestro sueño.

A mi angelito, Nico, mi gatito que estuvo conmigo desde el primer día, mi acompañante en las clases virtuales y por siempre esperarme en la ventana al regresar de clases. Aunque ya no estés conmigo, sigues siendo parte de esto, por ser mi motivo de desestrés, por tu gran compañía y amor.

AGRADECIMIENTO

En primer lugar, quiero agradecer a Dios, quien me ha otorgado la sabiduría para estar hoy culminando un sueño, por brindarme fortaleza para sobrellevar los momentos difíciles y por cuidar de mí y de mi familia.

Agradezco profundamente a mis padres, por su incansable guía, por su sacrificio y por permitirme cumplir una meta más en mi vida. Gracias por nutrirme de su sabiduría, por formar a la persona que soy hoy en día y por compartir conmigo sus valiosos valores. Gracias por ser los mejores padres y por confiar siempre en mí.

Mi sincero agradecimiento a todos los docentes que me han compartido su conocimiento y experiencia. Gracias por nutrirme con saberes que han sido clave en mi desarrollo académico y personal, semestre tras semestre.

Agradezco de manera especial a mi tutor, PhD. Francklin Rivas, y a mis lectores, PhD. Laura Guerra y PhD. Dulce Rivero, por su tiempo, dedicación y sabias recomendaciones, las cuales han sido esenciales para la culminación de este proyecto.

Gracias a la Pontificia Universidad Católica del Ecuador Ibarra, por brindarme la oportunidad de crecer tanto como persona como profesional, y por ser un espacio que me ha permitido desarrollar mis capacidades y conocimientos.

Finalmente, quiero agradecer a cada una de las personas que han sido parte de este proceso. A mis abuelos, a mi familia en general, mi más grande y profundo agradecimiento por su apoyo incondicional y por estar siempre a mi lado.

ÍNDICE

CERTIFICACIÓN TUTOR	ii
PÁGINA DE APROBACIÓN DEL TRIBUNAL	iii
ACTA DE CESIÓN DE DERECHOS	iv
AUTORIA	v
DEDICATORIA	vi
AGRADECIMIENTO	vii
ÍNDICE DE FIGURAS	xi
ÍNDICE DE TABLAS	xii
RESUMEN	xiii
ABSTRACT	xiv
INTRODUCCIÓN	1
CAPÍTULO I:	3
ESTADO DEL ARTE	3
<i>1.1 Estado del Arte</i>	<i>3</i>
1.1.1 Evolución del reconocimiento facial y análisis emocional	3
1.1.2 Era Pre-Digital (1960-1990).....	4
1.1.3 Era digital temprana (1990-2010)	5
1.1.4 Era del Deep Learning (2010-presente).....	5
<i>1.2 Fundamentos del análisis emocional</i>	<i>5</i>
1.2.1 Detección de puntos faciales	6
1.2.2 Análisis de expresiones	8
<i>1.3 Tecnologías actuales en reconocimiento facial</i>	<i>8</i>
1.3.1 Deep Learning y arquitecturas neuronales	8
1.3.2 Preprocesamiento y optimización.....	10
<i>1.4 Aplicaciones en entornos educativos</i>	<i>11</i>
1.4.1 Impacto en el bienestar docente.....	11
1.4.2 Integración con sistemas institucionales	11

1.5	<i>Impacto de las emociones en el entorno laboral educativo</i>	11
1.5.1	Inteligencia emocional artificial y su relevancia	12
1.6	<i>Tendencias actuales y direcciones futuras</i>	13
1.6.1	Inteligencia artificial ética	13
1.6.2	Innovaciones emergentes	13
1.7	<i>Desafíos y limitaciones actuales</i>	14
1.7.1	Precisión y sesgo	14
1.7.2	Dificultad en reconocer emociones complejas	14
1.7.3	Interpretación dependiente del contexto	14
1.7.4	Impacto psicológico del monitoreo	15
1.7.5	Redes neuronales convolucionales	15
CAPÍTULO II:		17
MATERIALES Y MÉTODOS		17
2.1	<i>Metodología de investigación</i>	17
2.2	<i>Metodología de desarrollo</i>	17
2.2.1	Fases del ciclo de vida del proyecto:	19
2.3	<i>Requisitos del sistema</i>	21
2.4	<i>Historias de usuario</i>	23
2.5	<i>Planificación de sprints</i>	31
2.6	<i>Preprocesamiento de imágenes</i>	32
2.7	<i>Diseño general del prototipo</i>	33
2.7.1	Funcionalidades de la interfaz	33
2.7.2	Funcionamiento del sistema de reconocimiento facial	34
2.7.3	Flujo de operación del prototipo	34
2.7.4	Eficiencia y automatización	35
2.7.5	Diagramas del Funcionamiento	35
2.7.6	Esquema de componentes del prototipo	36
2.8	<i>Entrenamiento del modelo de Deep Learning</i>	37
2.9	<i>Interfaz de usuario y retroalimentación en tiempo real</i>	41
2.9.1	Pantalla de espera	41
2.9.2	Interfaz de detección activa	42
2.10	<i>Tecnologías y herramientas utilizadas</i>	43
2.10.1	Hardware	43
2.11	<i>Lenguajes y bibliotecas de programación</i>	44

2.12 Dataset de entrenamiento: FER2013	45
2.13 Distribución de las imágenes	45
2.14 Validación del prototipo.....	46
CAPÍTULO III:	47
RESULTADOS Y DISCUSIÓN	47
3.1 Resultados del entrenamiento del modelo.....	47
3.1.1 Análisis detallado del tiempo de procesamiento	50
3.1.2 Factores que afectan el rendimiento	54
3.2 Resultados de la interfaz del prototipo.....	55
3.2.1 Interfaz de pantalla de espera	56
3.2.2 Interfaz de reconocimiento facial	57
3.2.3 Evaluación de usabilidad.....	58
3.3 Desarrollo y resultados del prototipo	59
3.3.1 Fase inicial del prototipo	59
3.3.2 Fase final del prototipo.....	61
3.4 Pruebas y evaluación del sistema.....	62
3.4.1 Pruebas de sistema en tiempo real.....	63
3.4.2 Validación del sistema.....	64
3.4.3 Evaluación del Hardware.....	70
3.4.4 Evaluación de Usabilidad	71
CONCLUSIONES	73
RECOMENDACIONES	74
REFERENCIAS BIBLIGRÁFICAS	75
ANEXOS	78

ÍNDICE DE FIGURAS

Figura 1 Representación de los 68 puntos faciales clave utilizados en el análisis de expresiones	7
Figura 2 Diagrama de flujo del sistema de reconocimiento facial y evaluación emocional	36
Figura 3 Esquema de componentes del prototipo de reconocimiento facial y evaluación emocional	37
Figura 4 Arquitectura del modelo de deep learning optimizado para clasificación de emociones	40
Figura 5 Boceto inicial de la pantalla de espera	41
Figura 6 Boceto de la interfaz durante la detección activa.....	42
Figura 7 Distribución de imágenes en el dataset FER2013.....	45
Figura 8 Precisión del modelo Mini-Xception por emoción y precisión general.....	48
Figura 9 Registro de tiempos de procesamiento en consola.....	51
Figura 10 Resumen estadístico de los tiempos de procesamiento.....	51
Figura 11 Gráfico de variabilidad de los tiempos de procesamiento	52
Figura 12 Histograma de distribución de tiempos de procesamiento.....	53
Figura 13 Pantalla de espera del sistema.....	56
Figura 14 Interfaz en acción mostrando video en vivo, emoción detectada y mensaje de retroalimentación	58
Figura 15 Prototipo durante la fase inicial de ensamblaje.....	60
Figura 16 Prototipo ensamblado en su diseño final, listo para implementación	62
Figura 17 Resultados de pruebas complementarias en un ambiente más iluminado	64
Figura 18 Matriz de confusión generada durante las pruebas con usuarios	66
Figura 19 Gráfico de barras que ilustra la precisión por emoción	67
Figura 20 Gráfico de barras que ilustra las emociones registradas por día	68
Figura 21 Gráfico de barras que ilustra las emociones registradas por mes	68
Figura 22 Docente interactuando con el sistema durante las pruebas	69

ÍNDICE DE TABLAS

Tabla 1 <i>Evolución del reconocimiento facial</i> _____	4
Tabla 2 <i>Comparativa de arquitecturas neuronales relevantes</i> _____	9
Tabla 3 Requisitos funcionales _____	21
Tabla 4 Requisitos no funcionales _____	22
Tabla 5 Historia de usuario para la configuración del entorno de desarrollo _____	23
Tabla 6 Historia de usuario para la identificación de docentes _____	24
Tabla 7 Historia de usuario para la detección facial _____	25
Tabla 8 Historia de usuario para la clasificación de emociones básicas _____	26
Tabla 9 Historia de Usuario para la Interfaz de Retroalimentación _____	27
Tabla 10 Historia de usuario para la activación y detección automática _____	28
Tabla 11 Historia de usuario para la evaluación y validación del Prototipo _____	29
Tabla 12 Historia de usuario para la optimización de rendimiento _____	30
Tabla 13 Resumen de sprint planificados _____	31
Tabla 14 Métricas de evaluación por emoción _____	49
Tabla 15 Métricas de rendimiento del sistema _____	63
Tabla 16 Métricas de precisión por emoción _____	65
Tabla 17 Resultados de la evaluación del rendimiento del hardware del prototipo _____	70
Tabla 18 Cumplimiento de requisitos del sistema _____	72

RESUMEN

El presente trabajo aborda el desarrollo de un prototipo de reconocimiento facial para la evaluación emocional del personal docente de la Escuela de Ingeniería de la Universidad Católica de Ecuador, sede Ibarra. Este sistema, diseñado para contribuir al bienestar de los docentes, emplea técnicas avanzadas de inteligencia artificial y visión por computadora para detectar y clasificar emociones en tiempo real.

Para su implementación, se utilizó el modelo mini-Xception, una arquitectura optimizada para tareas de reconocimiento facial y clasificación emocional en dispositivos de recursos limitados. El modelo fue entrenado con el conjunto de datos FER2013, compuesto por imágenes en escala de grises de rostros con diversas expresiones emocionales. La detección facial se realizó mediante OpenCV y Dlib, aplicando técnicas de ecualización del histograma y alineación facial para mejorar la precisión del reconocimiento. El entrenamiento del modelo fue optimizado con el algoritmo Adam, mientras que la evaluación del rendimiento se realizó con la función de pérdida CrossEntropyLoss.

El prototipo se activa automáticamente al detectar la presencia de una persona frente a la cámara, capturando y analizando su rostro para identificar emociones. Los resultados se presentan en tiempo real a través de una interfaz gráfica que muestra el estado emocional detectado, su nivel de confianza y otros detalles relevantes. Además, el sistema almacena los registros en formato JSON, incluyendo información como fecha, hora y emoción identificada, permitiendo un análisis histórico de los estados emocionales registrados.

La interfaz gráfica incorpora herramientas de visualización y análisis, como la exportación de datos y la generación de gráficos de barras. Estas funcionalidades permiten filtrar la información por períodos de tiempo, proporcionando una visión integral de los patrones emocionales. Este análisis puede ser utilizado por el personal administrativo para implementar estrategias que favorezcan el bienestar emocional de los docentes, permitiendo observar tendencias y patrones emocionales colectivos a lo largo del tiempo.

Palabras clave: Reconocimiento facial, evaluación emocional, mini-Xception, Pytorch, visión por computadora, inteligencia artificial.

ABSTRACT

This study focuses on the development of a facial recognition prototype designed to evaluate the emotions of faculty members at the School of Engineering of the Catholic University of Ecuador, Ibarra campus. This system, aimed at promoting faculty well-being, employs advanced artificial intelligence and computer vision techniques to detect and classify emotions in real time.

For its implementation, the mini-Xception model was utilized—an optimized architecture for facial recognition and emotional classification in resource-limited devices. The model was trained using the FER2013 dataset, which consists of grayscale images of faces displaying various emotional expressions.

Facial detection was performed using OpenCV and Dlib, applying histogram equalization and facial alignment techniques to enhance recognition accuracy. The model training process was optimized using the Adam algorithm, while performance evaluation was conducted using the CrossEntropyLoss function.

The prototype activates automatically upon detecting the presence of a person in front of the camera, capturing and analyzing their face to identify emotions. The results are displayed in real time through a graphical user interface that presents the detected emotional state, confidence level, and other relevant details. Additionally, the system stores records in JSON format, including information such as date, time, and identified emotion, allowing for historical analysis of emotional states.

The graphical interface also incorporates visualization and analysis tools, such as data export and bar graph generation. These features enable filtering information over time periods, providing a comprehensive view of emotional patterns. This analysis can be leveraged by administrative personnel to implement strategies that support faculty well-being, allowing the observation of collective emotional trends over time.

Keywords: Facial recognition, emotional evaluation, mini-Xception, Pytorch, computer vision, artificial intelligence

INTRODUCCIÓN

La evolución tecnológica ha impulsado el desarrollo de sistemas capaces de analizar y comprender el entorno humano mediante visión artificial, un campo de la inteligencia artificial que busca replicar las capacidades visuales humanas en las máquinas. Este avance ha permitido la creación de herramientas que no solo identifican rostros, sino que también pueden analizar estados emocionales, con aplicaciones en áreas como seguridad, salud mental y bienestar humano.

Estudios recientes han demostrado que el estado emocional de los docentes influye directamente en la calidad educativa, con investigaciones que indican que hasta un 63% del rendimiento académico puede verse afectado por el bienestar emocional del profesor. Un ambiente emocionalmente positivo en el aula no solo favorece el aprendizaje de los estudiantes, sino que también mejora la interacción entre docentes y alumnos, promoviendo un entorno educativo más eficiente y armonioso.

En la Pontificia Universidad Católica del Ecuador, sede Ibarra (PUCE-I), donde la interacción entre docentes y estudiantes es constante, el estado emocional del personal docente puede afectar significativamente su desempeño y la calidad del proceso educativo. Sin embargo, actualmente no existe un sistema que permita monitorear de manera objetiva las emociones de los docentes durante su registro diario en la institución. Esta ausencia de herramientas limita la posibilidad de implementar estrategias para el bienestar emocional del personal académico y, en consecuencia, su impacto en el aprendizaje estudiantil.

Ante esta problemática, el presente trabajo propone el desarrollo de un prototipo de reconocimiento facial orientado a la evaluación emocional de los docentes en el entorno académico. Este prototipo implementa un sistema de reconocimiento que tiene como propósito proporcionar retroalimentación en tiempo real sobre el estado emocional del personal docente, así como herramientas de análisis y visualización de datos que permitan identificar patrones emocionales a lo largo del tiempo. De esta manera, se busca contribuir a la generación de un entorno laboral más positivo, que fortalezca el bienestar docente y optimice la calidad educativa.

La falta de herramientas para identificar, analizar y gestionar las emociones dentro del ámbito educativo constituye una brecha que este proyecto pretende abordar. Al integrar inteligencia artificial y visión por computadora en un sistema automatizado, se espera que esta solución represente un paso significativo en la incorporación de tecnologías emergentes para el monitoreo y análisis de la salud emocional en contextos educativos.

Objetivos

Objetivo General

Desarrollar un prototipo de reconocimiento facial para la evaluación emocional del personal docente de la Escuela de Ingeniería de la Universidad Católica del Ecuador, sede Ibarra, con el fin de mejorar el ambiente laboral y la calidad de la enseñanza.

Objetivos Específicos

- Desarrollar un algoritmo de detección y clasificación de emociones basado en las expresiones faciales capturadas.
- Diseñar una interfaz de usuario que proporcione retroalimentación emocional en tiempo real al personal docente.
- Evaluar la eficacia y precisión del prototipo en condiciones reales de uso en la Escuela de Ingeniería de la PUCEI.

Alcance y limitaciones del estudio

Este estudio se centró en el desarrollo de un prototipo que implementa un sistema de reconocimiento facial para evaluar las emociones básicas (alegría, tristeza, enojo, sorpresa) en tiempo real, aplicable inicialmente a un grupo seleccionado de docentes de la Escuela de Ingeniería de la PUCE-I. Las limitaciones del estudio incluyen las restricciones de hardware y software para el procesamiento de imágenes en tiempo real en dispositivos de bajo costo, como la Raspberry Pi, y el enfoque en un número reducido de emociones, sin considerar emociones más complejas o combinadas.

CAPÍTULO I:

ESTADO DEL ARTE

Este capítulo presenta una revisión bibliográfica sobre investigaciones y avances relacionados con el reconocimiento facial y el análisis emocional. Se abordan los fundamentos conceptuales, la evolución tecnológica y las aplicaciones actuales en entornos educativos. Además, se identifican las metodologías, herramientas y desafíos predominantes, proporcionando una base sólida para el desarrollo del presente trabajo.

1.1 Estado del Arte

El reconocimiento facial y el análisis emocional han experimentado un crecimiento significativo en los últimos años, gracias a los avances en inteligencia artificial y visión por computadora. La evolución de esta tecnología ha sido notable: mientras que los primeros sistemas desarrollados por Bledsoe, Wolf y Bisson en la década de 1960 dependían del marcado manual de puntos faciales, los sistemas actuales utilizan redes neuronales profundas que automatizan completamente este proceso. Como señalan Li et al. (2020) en su investigación sobre sistemas de reconocimiento facial en tiempo real, las arquitecturas modernas basadas en *deep learning* pueden detectar y analizar más de 68 puntos faciales automáticamente, con una precisión superior al 99% en condiciones controladas.

Según Zhang y Liu (2022), el análisis emocional mediante visión por computadora ha revolucionado diversos campos, desde la seguridad hasta la educación. En el contexto educativo específicamente, la implementación de tecnologías de monitoreo emocional ha comenzado a demostrar un impacto positivo, no solo en el bienestar del personal docente, sino también en la calidad de la enseñanza. Este capítulo ofrece una revisión crítica de las investigaciones más relevantes en el área del reconocimiento facial y análisis emocional, con un enfoque en su aplicación en entornos educativos.

1.1.1 Evolución del reconocimiento facial y análisis emocional

El reconocimiento facial ha experimentado una evolución notable desde su surgimiento en la década de 1960, marcada por avances tecnológicos que han transformado

su precisión y aplicabilidad. Esta evolución se puede dividir en tres eras, destacando su impacto en la inteligencia artificial y el *machine learning*.

Las tres eras clave son: la era pre-digital, la era digital temprana y la era del *deep learning*. Cada una de estas fases posee hitos importantes que han moldeado el desarrollo de los sistemas actuales. En la Tabla 1, se presentan las características distintivas y logros significativos de cada etapa.

Tabla 1 Evolución del reconocimiento facial

Era	Periodo	Características	Hitos principales
Pre-Digital	1960-1990	Métodos geométricos básicos, análisis manual.	Desarrollo de puntos de referencia facial (Li y Jain, 2018).
Digital Temprana	1990-2010	Algoritmos automatizados, inicio del uso de PCA y Eigenfaces.	Desarrollo del método Eigenfaces (Sirovich, 1987).
Deep Learning	2010-presente	Redes neuronales convolucionales, alta precisión, procesamiento en tiempo real.	DeepFace alcanza 97.35% de precisión; FaceNet logra 99.63% en LFW dataset (Guo et al., 2021).

1.1.2 Era Pre-Digital (1960-1990)

Durante esta era, los métodos geométricos básicos dominaban el campo del reconocimiento facial. Los investigadores dependían de análisis manuales de características faciales, lo que resultaba en una precisión limitada y un procesamiento lento. Aunque estas

primeras aproximaciones sentaron las bases del reconocimiento facial, la tecnología estaba muy restringida por la capacidad computacional de la época.

1.1.3 Era digital temprana (1990-2010)

Con el advenimiento de algoritmos computacionales, se introdujeron los primeros sistemas automatizados de reconocimiento facial. Durante este periodo, las mejoras en la capacidad de procesamiento permitieron el desarrollo de enfoques más sofisticados, que comenzaron a abordar algunas de las limitaciones de las técnicas anteriores.

1.1.4 Era del Deep Learning (2010-presente)

La introducción de redes neuronales convolucionales (CNNs) marcó un punto de inflexión en el reconocimiento facial. Guo et al. (2021) documentan cómo estas arquitecturas modernas han alcanzado niveles de precisión excepcionales, con FaceNet logrando un 99.63% de precisión en el conjunto de datos Labeled Faces in the Wild (LFW). Wang y Deng (2021) señalan que el uso de CNNs ha superado las limitaciones de los modelos tradicionales basados en características geométricas, lo que ha transformado la eficacia del reconocimiento facial en diversas aplicaciones. En cuanto al análisis emocional, el uso de *deep learning* ha permitido la automatización del reconocimiento de expresiones faciales. La combinación de algoritmos avanzados con el aumento en la capacidad de procesamiento ha facilitado la detección de emociones en tiempo real, integrando estos sistemas en aplicaciones prácticas como el monitoreo emocional en entornos laborales y educativos.

1.2 Fundamentos del análisis emocional

El análisis emocional mediante reconocimiento facial se fundamenta en la teoría de las emociones básicas propuesta por Ekman (2003), quien identificó seis emociones universales: alegría, tristeza, ira, miedo, sorpresa y asco. Según este autor, dichas emociones son reconocidas de manera casi universal a través de expresiones faciales, independientemente de la cultura o el contexto, lo que las convierte en un marco robusto para el análisis emocional.

El reconocimiento emocional puede definirse como la capacidad de identificar con precisión los estados emocionales de las personas a partir de diferentes señales, como expresiones faciales, lenguaje corporal y tono de voz. Khare et al. (2024) describen las emociones como estados dinámicos que surgen de interacciones complejas entre el individuo y su entorno, influyendo significativamente en el comportamiento y bienestar de las personas. Estas respuestas emocionales reflejan tanto aspectos cognitivos como fisiológicos, lo que resalta su naturaleza multifacética. Existen dos enfoques principales para comprender las emociones:

Modelo discreto: Propuesto por psicólogos como Paul Ekman, este modelo clasifica las emociones en categorías universales y fijas. Las seis emociones básicas identificadas (felicidad, tristeza, ira, sorpresa, miedo y disgusto) se asocian con patrones faciales específicos que pueden reconocerse universalmente, independientemente del contexto cultural.

Modelo multidimensional: Este modelo, representado por autores como James Russell, plantea que las emociones son experiencias complejas e influenciadas por factores personales, culturales y sociales. En lugar de clasificar emociones específicas, se enfoca en dimensiones emocionales clave que definen el estado emocional de una persona:

Valencia: Define si la emoción es positiva o negativa.

Arousal (Activación): Indica el nivel de energía asociado con la emoción.

Dominancia: Describe la sensación de control o de ser dominado por la emoción.

El modelo multidimensional permite una comprensión más detallada de los estados emocionales, al considerar que las emociones no son experiencias estáticas, sino dinámicas y contextuales. Este enfoque amplía la perspectiva al incorporar la diversidad y complejidad inherente a las emociones humanas.

1.2.1 Detección de puntos faciales

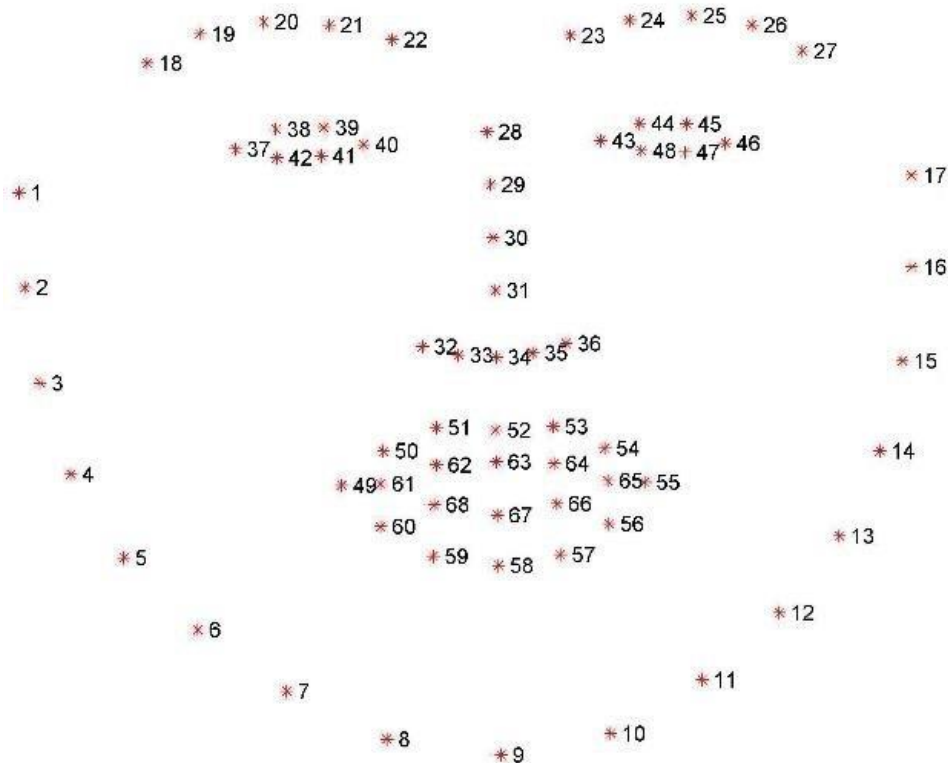
La detección de puntos faciales es una técnica esencial en el análisis de expresiones faciales, utilizada ampliamente en sistemas de visión por computadora y reconocimiento

facial. Este método permite identificar 68 puntos clave, conocidos como *landmarks*, distribuidos de manera estratégica en áreas específicas del rostro, tales como los ojos, la nariz, la boca, las cejas y la mandíbula (Ferrin et al., 2020). Estos puntos actúan como referencias geométricas precisas que permiten una interpretación más detallada de las expresiones y movimientos faciales, mejorando la capacidad del sistema para reconocer emociones, gestos y patrones de comportamiento.

Además de facilitar la interpretación de las expresiones faciales, la detección de estos puntos es fundamental para la alineación y normalización de las imágenes. Esto asegura que el análisis sea consistente y preciso incluso en condiciones variables, como cambios en los ángulos de captura, iluminación diferente o variaciones en las características faciales entre diferentes personas.

Como se ilustra en la Figura 1, estos 68 puntos faciales son fundamentales para asegurar un análisis confiable y preciso en aplicaciones de visión por computadora, desde el reconocimiento de emociones hasta la identificación de individuos en entornos controlados.

Figura 1 Representación de los 68 puntos faciales clave utilizados en el análisis de expresiones



Los puntos faciales se distribuyen estratégicamente sobre las principales características del rostro, como los ojos, nariz, boca y mandíbula. Estas referencias geométricas son fundamentales para el análisis y clasificación de emociones en sistemas de visión artificial. Fuente: Ferrin, Carlos, Mosquera De La Cruz, José y Loaiza-Correa, Humberto (2020).

1.2.2 Análisis de expresiones

Los sistemas de reconocimiento emocional también se centran en la clasificación de micro expresiones y la interpretación de patrones musculares, evaluando la intensidad emocional de las expresiones detectadas.

1.3 Tecnologías actuales en reconocimiento facial

1.3.1 Deep Learning y arquitecturas neuronales

El *deep learning* representa una rama fundamental del aprendizaje automático que utiliza redes neuronales con múltiples capas para el procesamiento de información. En el contexto del reconocimiento facial, diversas arquitecturas neuronales han sido desarrolladas para abordar diferentes aspectos del procesamiento de imágenes faciales.

Una de estas arquitecturas es el mini-Xception, que surge como una adaptación optimizada de la arquitectura Xception original (Chollet, 2017). Este modelo se caracteriza por su estructura de capas convolucionales separables en profundidad, lo que permite reducir significativamente el número de parámetros mientras mantiene un alto rendimiento en tareas de clasificación.

FaceNet, introducido por Schroff et al. (2015), representa un enfoque innovador en el reconocimiento facial al introducir el concepto de embeddings faciales. Esta arquitectura utiliza una red neuronal profunda para generar representaciones vectoriales de 128 dimensiones de los rostros, permitiendo una comparación directa entre diferentes imágenes faciales mediante distancias euclidianas.

DeepFace (Taigman et al., 2014), por su parte, introduce un método de alineación facial en 3D como paso previo al procesamiento, seguido de una red neuronal convolucional

profunda. Esta arquitectura implementa un proceso de normalización facial que permite manejar variaciones en la pose y la iluminación.

MobileNetV2 (Sandler et al., 2018) se distingue por su arquitectura optimizada para dispositivos con recursos limitados. Utiliza bloques residuales invertidos y convoluciones separables en profundidad, lo que permite reducir significativamente el costo computacional mientras mantiene un buen rendimiento en tareas de reconocimiento facial.

ResNet-50 (He et al., 2016) es una arquitectura que implementa conexiones residuales para facilitar el entrenamiento de redes muy profundas. Este modelo utiliza bloques residuales que permiten que la información se propague mejor a través de la red, lo que resulta en una mayor capacidad para aprender características complejas de las imágenes faciales.

En el ámbito educativo, estas tecnologías han demostrado ser herramientas efectivas para analizar emociones y patrones faciales, facilitando nuevas formas de interacción y monitoreo emocional. Esto no solo mejora la experiencia de aprendizaje, sino también el bienestar emocional de los docentes y estudiantes. La Tabla 2 presenta una comparación detallada de las arquitecturas neuronales más relevantes utilizadas en el reconocimiento facial, resaltando sus ventajas y limitaciones, particularmente en el contexto de aplicaciones educativas.

Tabla 2 *Comparativa de arquitecturas neuronales relevantes*

Arquitectura	Precisión	Tiempo de Inferencia	Tamaño del Modelo	Ventajas	Limitaciones
Mini-Xception	95%	<100ms	2.5MB	Bajo consumo de recursos Buena precisión en tiempo real	Limitado a emociones básicas Requiere buena iluminación
FaceNet	99.63%	~200ms	30MB	Alta precisión	Mayor consumo de recursos

				Embedding facial robusto Excelente para verificación	Requiere GPU para rendimiento óptimo
DeepFace	97.35%	~150ms	25MB	Precisión cercana a humanos. Robusto a variaciones Bueno en poses diferentes	Computacionalmente intensivo Requiere alineación facial
MobileNetV2	92%	<50ms	3MB	Ultra liviano Muy rápido Ideal para móviles	Menor precisión Sensible a condiciones adversas
ResNet-50	98%	~180ms	98MB	Muy preciso Robusto Bien documentado	Pesado Alto consumo de recursos

1.3.2 Preprocesamiento y optimización

Para mejorar la precisión en la detección facial y emocional, se utilizan técnicas avanzadas de preprocesamiento, que preparan las imágenes antes de ser procesadas por el modelo de *deep learning*. Entre las técnicas más destacadas se encuentran:

- **Ecualización adaptativa del histograma:** Mejora la calidad de las imágenes ajustando el contraste, lo que es especialmente útil en situaciones de iluminación variable.

- **Detección de puntos faciales con Dlib:** Herramienta utilizada para la alineación automática de las características faciales (ojos, nariz, boca).
- **Alineación facial automática:** Garantiza que los rostros se encuentren en una posición estándar antes de ser procesados, mejorando la robustez del sistema.

1.4 Aplicaciones en entornos educativos

1.4.1 Impacto en el bienestar docente

En el ámbito educativo, la implementación de sistemas de monitoreo emocional ha demostrado un impacto significativo en la calidad de la enseñanza y el bienestar de los docentes. Liu y Chen (2022) encontraron que el uso de estas tecnologías en instituciones educativas mejoró la satisfacción laboral en un 15% y redujo el estrés en un 20%. Además, los docentes reportaron una mayor participación estudiantil cuando sus emociones eran monitoreadas y gestionadas adecuadamente.

1.4.2 Integración con sistemas institucionales

La implementación de estas tecnologías se realiza mediante sistemas de gestión que permiten el monitoreo de la asistencia, la evaluación continua y el análisis de desempeño. Las plataformas de recursos humanos pueden utilizar el reconocimiento emocional para gestionar el bienestar de los docentes, ofreciendo programas de apoyo personalizado y seguimiento de indicadores emocionales.

1.5 Impacto de las emociones en el entorno laboral educativo

El monitoreo emocional en el entorno educativo no solo busca mejorar la calidad de la enseñanza, sino también el bienestar emocional de los docentes. Según Seo, Barrett y Bartunek (2004), las emociones desempeñan un papel crucial en la motivación de los empleados, influyendo directamente en su satisfacción y productividad. En un contexto educativo, esta relación se traduce en la mejora del rendimiento docente y en la creación de un ambiente de trabajo más positivo.

Los estudios de Mora (2013), citados por Nieto (2020), subrayan la importancia fundamental de las emociones en los procesos de aprendizaje y memoria, estableciendo una

conexión directa entre la gestión emocional efectiva y el desempeño laboral óptimo. Esta perspectiva refuerza la necesidad de desarrollar sistemas robustos que faciliten el monitoreo y la gestión efectiva de las emociones del personal docente.

Además, Danvila y Sastre (2010) han aportado investigaciones valiosas que demuestran cómo las emociones influyen en la satisfacción y productividad de los empleados, reforzando la idea de que el bienestar emocional es un factor clave en el éxito organizacional. Páez y da Costa (2014) señalaron que la regulación emocional se ha estudiado tradicionalmente desde una perspectiva intraindividual, pero influir en las emociones de otros es crucial en la vida social. Un estudio realizado en una prisión terapéutica reveló que el 62% de los presos reportó una influencia emocional positiva, al igual que el 90% de los trabajadores.

1.5.1 Inteligencia emocional artificial y su relevancia

La Inteligencia Emocional Artificial (*Emotion AI*), también conocida como computación afectiva, es un campo emergente que busca procesar y replicar emociones humanas para mejorar las interacciones entre humanos y máquinas. Winfield (2020) reflexiona sobre la complejidad de este campo señalando: "Ahora hemos aceptado después de 60 años de IA que las cosas que originalmente pensamos que eran fáciles son en realidad muy difíciles". Esto refleja los retos inherentes al desarrollo tecnológico para captar emociones humanas auténticas.

Hernández y Picard (2022) del MIT Media Lab plantean una pregunta fundamental para el campo: "¿Cómo puede una máquina comunicar efectivamente si no sabe tu estado emocional?". Esta interrogante resalta la importancia crucial del contexto emocional para lograr interacciones significativas.

Un ejemplo notable en este campo es Affective, empresa fundada por el Kaliouby y Picard (2019), que ha desarrollado lo que denominan "Inteligencia Emocional Multimodal". Esta tecnología analiza expresiones faciales y lenguaje corporal durante interacciones publicitarias, logrando niveles precisos gracias a un conjunto diverso de datos utilizado para entrenar sus algoritmos.

1.6 Tendencias actuales y direcciones futuras

1.6.1 Inteligencia artificial ética

El desarrollo de sistemas de reconocimiento facial y análisis emocional presenta varios desafíos éticos. Es fundamental garantizar que estos sistemas respeten la privacidad de los usuarios y se utilicen de manera justa y responsable. Entre las principales consideraciones éticas se encuentran:

- **Privacidad de los datos biométricos:** La protección de los datos sensibles capturados por los sistemas es un aspecto clave.
- **Consentimiento informado:** Los usuarios deben estar plenamente informados y dar su consentimiento antes de que sus datos sean utilizados.
- **Seguridad de la información:** Los sistemas deben estar diseñados para prevenir accesos no autorizados y garantizar que la información biométrica esté protegida.
- **Uso responsable de la IA:** La inteligencia artificial debe ser utilizada con fines que respeten los derechos de los individuos y no perpetúen sesgos o discriminación.

1.6.2 Innovaciones emergentes

Las tendencias actuales en el campo del reconocimiento facial y el análisis emocional incluyen el desarrollo de tecnologías que combinan múltiples fuentes de datos para mejorar la precisión y confiabilidad de los sistemas. Los sistemas multimodales, que integran el reconocimiento facial con análisis de voz, han demostrado mejoras significativas en la precisión de la detección emocional. Los algoritmos de aprendizaje continuo están revolucionando la capacidad de los sistemas para adaptarse y mejorar constantemente a medida que procesan nuevos datos.

La integración con tecnologías de realidad aumentada está abriendo nuevas posibilidades para proporcionar retroalimentación emocional en tiempo real, mejorando significativamente la interacción humano-computadora. Los sistemas de retroalimentación instantánea están permitiendo respuestas más inmediatas y precisas en diversas aplicaciones educativas y profesionales.

1.7 Desafíos y limitaciones actuales

A pesar de los avances significativos en el campo, persisten diversos desafíos técnicos que requieren atención continua.

1.7.1 Precisión y sesgo

Estudios como el de Buolamwini y Gebru (2018) han demostrado que la precisión del reconocimiento emocional varía significativamente entre diferentes grupos demográficos. Estos sesgos son el resultado de la falta de diversidad en los conjuntos de datos, lo cual afecta la precisión y equidad de los sistemas de reconocimiento facial. La incorporación de conjuntos de datos más diversos y representativos es crucial para garantizar la precisión y evitar la discriminación de género, raza o etnia en estos sistemas.

1.7.2 Dificultad en reconocer emociones complejas

Los sistemas actuales suelen clasificar las emociones en categorías básicas, lo que limita la capacidad de reconocer emociones más complejas, como la ambivalencia. Según Mollahosseini et al. (2017), los modelos de reconocimiento emocional reducen las expresiones a patrones simples, lo que ocasiona que no se capturen los matices emocionales críticos, especialmente en contextos educativos donde las emociones pueden ser más complejas y sutiles.

1.7.3 Interpretación dependiente del contexto

El reconocimiento de emociones también depende de factores contextuales y culturales que los sistemas de inteligencia artificial actuales no siempre interpretan correctamente. Kätsyri et al. (2015) explican cómo el contexto social influye en la percepción emocional, destacando la necesidad de que los modelos de reconocimiento emocional mejoren en este sentido para evitar interpretaciones erróneas. El mismo gesto o expresión facial puede tener diferentes significados dependiendo del entorno y la cultura, lo que representa un desafío adicional para los sistemas automatizados.

1.7.4 Impacto psicológico del monitoreo

El monitoreo constante del estado emocional puede generar ansiedad y estrés en los individuos. Nissenbaum (2010) argumenta que los sistemas de reconocimiento emocional deben implementarse con precaución, especialmente en entornos educativos, para minimizar los efectos psicológicos negativos sobre los docentes y estudiantes. La vigilancia emocional constante puede tener implicaciones éticas y psicológicas, que deben ser cuidadosamente consideradas en el diseño de este tipo de sistemas.

El reconocimiento facial y el análisis emocional han experimentado un notable crecimiento gracias a los avances en *deep learning* y tecnologías de visión artificial. La integración de estas herramientas en entornos educativos ofrece una oportunidad única para mejorar el bienestar de los docentes y, en última instancia, la calidad de la enseñanza. Sin embargo, para una implementación exitosa y responsable, es necesario abordar diversos desafíos técnicos, prácticos y éticos. Este proyecto tiene como objetivo contribuir a este campo mediante el desarrollo de un sistema de reconocimiento emocional que combine las mejores prácticas tecnológicas con un enfoque en la mejora del ambiente laboral en instituciones educativas.

1.7.5 Redes neuronales convolucionales

Las Redes Neuronales Convolucionales (CNN) son un tipo especializado de redes neuronales profundas, diseñadas para procesar datos con estructura de cuadrícula, como imágenes. Inspiradas en la organización del córtex visual biológico, estas redes han revolucionado áreas como el reconocimiento facial, la clasificación de imágenes y el análisis emocional debido a su capacidad para extraer características automáticamente y su alto rendimiento en tareas de visión por computadora (LeCun et al., 2015).

Arquitectura de las redes neuronales convolucionales

Una CNN típica está compuesta por tres tipos principales de capas, cada una con una función específica en el procesamiento y análisis de las imágenes (Goodfellow et al., 2016):

- **Capas convolucionales:** Estas capas funcionan como detectores de características aplicando filtros (kernels) que recorren la imagen y extraen

patrones locales, como bordes, texturas y formas. En el contexto del reconocimiento emocional, estas capas aprenden a identificar características relevantes de las expresiones faciales, como la curvatura de los labios, el levantamiento de las cejas o el ensanchamiento de los ojos.

- **Capas de pooling:** También conocidas como capas de submuestreo, estas reducen la dimensionalidad de los datos mientras mantienen la información más relevante. Esto ayuda a disminuir la carga computacional y hace que el modelo sea más robusto frente a pequeñas variaciones en las características faciales, como desplazamientos o cambios en el ángulo de la cabeza.
- **Capas completamente conectadas:** Situadas al final de la arquitectura, estas capas combinan todas las características extraídas y las utilizan para realizar la clasificación final. En el caso del reconocimiento emocional, esta capa determina la emoción detectada entre las categorías predefinidas, como felicidad, tristeza, enojo o sorpresa.

Funcionamiento en el reconocimiento emocional

- **Procesamiento inicial:** La imagen facial, normalmente preprocesada (escalada a un tamaño uniforme, en escala de grises y normalizada), se introduce en la red.
- **Extracción de características:** Las capas convolucionales analizan la imagen, detectando características faciales clave relacionadas con expresiones emocionales.
- **Reducción de dimensionalidad:** Las capas de pooling condensan la información para eliminar redundancias y resaltar las características más importantes.
- **Clasificación:** Las capas completamente conectadas combinan las características extraídas y clasifican la emoción en una de las categorías definidas.

Las CNNs son ampliamente utilizadas en el análisis emocional debido a su capacidad para detectar patrones complejos en imágenes faciales, incluso en condiciones desafiantes como cambios de iluminación o expresiones faciales sutiles (Zhang et al., 2023). Su capacidad para aprender características automáticamente elimina la necesidad de ingeniería manual, y su arquitectura optimizada permite obtener resultados precisos y eficientes.

CAPÍTULO II:

MATERIALES Y MÉTODOS

En este capítulo se describe de manera detallada la metodología de investigación y desarrollo adoptada para la implementación del prototipo de reconocimiento facial y evaluación emocional. El enfoque metodológico se basa en una combinación de métodos científicos aplicados y técnicas de desarrollo adaptadas a proyectos de inteligencia artificial. El objetivo es garantizar que el prototipo sea funcional, eficiente y adecuado al entorno educativo en el que se va a implementar. Para ello, se abordan tanto los métodos de investigación utilizados para evaluar el estado emocional de los docentes como las estrategias de desarrollo empleadas para crear un prototipo robusto. Este capítulo también presenta las fases del ciclo de vida del desarrollo, desde la definición de requisitos hasta la evaluación final.

2.1 Metodología de investigación

El proyecto se desarrolló como un estudio aplicado con enfoque cuantitativo, ya que se centra en la medición y análisis de datos numéricos específicos. Para evaluar el rendimiento del sistema, se midieron los siguientes parámetros:

- Porcentajes de precisión en la detección de emociones
- Tiempos de respuesta del prototipo medidos en milisegundos
- Tasas de error en el reconocimiento facial
- Métricas de rendimiento del hardware

2.2 Metodología de desarrollo

Para el desarrollo del sistema de reconocimiento facial y evaluación emocional, se adoptó una metodología ágil, específicamente un enfoque iterativo adaptado a proyectos de inteligencia artificial. Dado que no era posible integrar directamente el sistema con el reloj biométrico de la institución, se decidió implementar el reconocimiento facial en un entorno de desarrollo local. Este enfoque permitió probar el sistema en un entorno controlado y accesible, lo que facilitó su validación antes de la integración al prototipo final. La

flexibilidad proporcionada por marcos como SCRUM se combinó con un ciclo de vida especializado en sistemas de *machine learning*, lo que permitió una construcción continua del sistema, con ajustes en el enfoque según los resultados obtenidos en cada fase del proyecto. La implementación en este entorno de pruebas permitió validar el sistema de reconocimiento facial en condiciones similares a las del entorno real, evaluando su precisión y rendimiento antes de su integración al dispositivo biométrico.

El sistema implementa dos interfaces principales que gestionan diferentes estados de operación:

- **Pantalla de espera:** Se activa cuando el sistema está en reposo, mostrando el logo del proyecto, la hora actual y un mensaje indicando que el sistema está listo para la detección.
- **Interfaz de reconocimiento facial:** Se activa cuando un usuario se encuentra frente a la cámara, permitiendo la captura y análisis de la emoción detectada, mostrando la retroalimentación visual en tiempo real.

La metodología de investigación empleada integró tres métodos complementarios que permitieron realizar un análisis exhaustivo y validar el sistema de manera efectiva. Estos métodos fueron seleccionados por su capacidad para abordar tanto los aspectos cualitativos como cuantitativos del desarrollo del sistema, asegurando un enfoque integral que combina la identificación de necesidades, el análisis técnico y la verificación práctica.

Método descriptivo:

Este enfoque facilitó la identificación y caracterización de las emociones predominantes en el personal docente, proporcionando una visión precisa de su estado emocional en el entorno laboral. Esta información fue esencial para definir los objetivos del sistema y adaptar su diseño a las necesidades reales del contexto educativo.

Método analítico:

A través de este método, se desglosaron y examinaron los componentes tecnológicos involucrados en la implementación del sistema. Se estudiaron en detalle los algoritmos de

reconocimiento facial, las técnicas de preprocesamiento de imágenes y la interacción con dispositivos de hardware como cámaras y pantallas. Este análisis garantizó que los procesos tecnológicos fueran confiables, consistentes y capaces de operar eficazmente en condiciones variables, como las que se encuentran en el entorno educativo.

Método experimental:

El método experimental permitió validar el prototipo desarrollado a través de pruebas controladas. Se evaluaron aspectos clave como la precisión del modelo de reconocimiento emocional, el tiempo de respuesta del sistema y la experiencia del usuario final. Las pruebas se realizaron tanto en condiciones simuladas como en escenarios reales, asegurando la fiabilidad y aplicabilidad del sistema en contextos educativos.

Durante las fases de diseño de la arquitectura y construcción, se definió un flujo de operación claro para el sistema. Este flujo abarcó funciones clave como la activación, captura, procesamiento, análisis y retroalimentación, asegurando que todas las etapas operaran de manera eficiente y coordinada. Este diseño facilitó la realización de pruebas de integración y permitió validar que cada componente cumpliera con los requisitos funcionales y no funcionales establecidos.

El ciclo de vida del desarrollo se estructuró en diversas fases clave, desde la definición inicial de los requisitos hasta el mantenimiento continuo del sistema. Este enfoque garantizó un desarrollo sistemático del proyecto, atendiendo tanto a los aspectos técnicos como a las necesidades específicas del contexto educativo donde se implementará el sistema. La metodología empleada permitió integrar eficazmente las diferentes etapas del desarrollo y realizar ajustes según fuera necesario, asegurando que el producto final fuera adecuado para su implementación en la institución educativa.

2.2.1 Fases del ciclo de vida del proyecto:

Definición de requisitos

Durante esta fase inicial, se identificaron las necesidades técnicas y funcionales del prototipo. Los requisitos se centraron en garantizar que el prototipo fuera capaz de detectar

emociones con alta precisión, integrarse con el entorno educativo y ofrecer retroalimentación emocional efectiva a los docentes. Este análisis de requisitos fue fundamental para el desarrollo del prototipo.

Diseño de arquitectura

En esta etapa se seleccionaron las herramientas, bibliotecas y arquitecturas de redes neuronales necesarias para implementar el sistema de reconocimiento facial. Se decidió usar el modelo *mini-Xception* debido a su eficiencia en el reconocimiento facial y clasificación de emociones en tiempo real, especialmente adecuado para dispositivos con recursos limitados como la Raspberry Pi.

Recolección y preparación de datos

Esta fase fue crucial para garantizar que el *dataset* utilizado fuera de alta calidad y estuviera adecuadamente preprocesado. Se implementaron técnicas avanzadas, como la normalización de las imágenes faciales y la alineación precisa de los rostros, con el fin de garantizar que los datos estuvieran listos para el análisis del modelo. Este proceso de preparación fue fundamental para obtener resultados consistentes y confiables en el sistema de reconocimiento emocional, al reducir las posibles variaciones y ruidos en las imágenes que podrían afectar la precisión del modelo.

Entrenamiento del modelo

Durante el entrenamiento, se ajustaron los hiperparámetros del modelo y se utilizó el conjunto de datos FER2013. Se aplicó la técnica de validación cruzada para asegurar la robustez del modelo y minimizar el riesgo de sobreajuste, garantizando así que el sistema fuera capaz de generalizar bien en condiciones reales.

Construcción del sistema

En esta fase, se desarrolló una interfaz que permite visualizar los resultados del sistema de reconocimiento facial. El prototipo integra los componentes de hardware y software necesarios para su funcionamiento en el entorno educativo.

Pruebas

Las pruebas se realizaron tanto en entornos controlados como en situaciones reales. Se validaron aspectos como la precisión del modelo, la estabilidad del sistema y la experiencia del usuario final. Las pruebas en el entorno educativo fueron especialmente importantes para asegurar que el sistema fuera adecuado para el contexto de trabajo de los docentes.

2.3 Requisitos del sistema

El sistema está diseñado para cumplir con una serie de requisitos funcionales y no funcionales que garantizan su operatividad y rendimiento dentro del entorno educativo.

Requisitos funcionales:

Los requisitos funcionales (Tabla 3), describen las principales funciones que el prototipo debe realizar para cumplir con los objetivos del proyecto. Estos incluyen la implementación del sistema de reconocimiento facial para la detección de emociones en tiempo real y la generación de retroalimentación visual.

Tabla 3 *Requisitos funcionales*

REQUISITOS FUNCIONALES	
Código	Descripción
RF-01	El sistema debe permanecer inactivo hasta que un docente se acerque al prototipo.
RF-02	El sistema debe activar la detección facial al detectar la proximidad del docente.
RF-03	El sistema debe capturar una imagen del rostro del docente y procesarla en tiempo real para detectar y clasificar la emoción.

RF-04 El sistema debe mostrar en la pantalla de 7 pulgadas la emoción detectada y proporcionar simultáneamente un mensaje de retroalimentación personalizado basado en dicha emoción.

RF-05 El sistema debe operar como un módulo independiente que capture y procese las emociones, almacene los registros de emociones detectadas en un archivo JSON y proporcione una interfaz de comunicación estándar para futura integración con el reloj biométrico.

Requisitos no funcionales:

Los requisitos no funcionales (Tabla 4), describen las características que el sistema debe cumplir para garantizar su calidad, seguridad, rendimiento y escalabilidad en el futuro.

Tabla 4 *Requisitos no funcionales*

REQUISITOS NO FUNCIONALES	
Categoría	Descripción
Usabilidad	El sistema debe ser fácil de usar por los docentes y configurar por el administrador técnico en menos de 2 horas.
Seguridad	El sistema debe garantizar la privacidad de los datos de emociones y accesibilidad limitada a usuarios autorizados.
Eficiencia	El sistema debe procesar y mostrar la emoción detectada en menos de 3 segundos.
Interoperabilidad	El sistema debe ser compatible con diferentes cámaras y sensores de proximidad para futuras actualizaciones.

2.4 Historias de usuario

Las historias de usuario son fundamentales para la construcción de sistemas ágiles. Estas historias describen las necesidades y expectativas de los usuarios del sistema, proporcionando una base clara para el desarrollo y asegurando que las funcionalidades del sistema estén alineadas con lo que los usuarios necesitan.

Configuración del entorno de desarrollo

La Tabla 5 presenta la historia de usuario número 1, denominada "Configuración del Entorno de Desarrollo", que tiene una alta prioridad tanto para el negocio como para el desarrollo. Esta historia describe la necesidad del usuario técnico de contar con un entorno de trabajo completamente configurado para el desarrollo del prototipo de reconocimiento facial y evaluación emocional. El entorno debe incluir las herramientas necesarias como Anaconda, Python, PyTorch, OpenCV, y dlib para asegurar un desarrollo eficiente. Los criterios de aceptación establecen los requisitos fundamentales que garantizan un entorno funcional y listo para el desarrollo, incluyendo la verificación de compatibilidad entre herramientas y la capacidad de realizar pruebas preliminares antes de la implementación en hardware.

Tabla 5 Historia de usuario para la configuración del entorno de desarrollo

Historia de Usuario	Número: HU-01	Nombre: Configuración del entorno de desarrollo
Usuario:	Usuario técnico	
Prioridad en negocio: Alta	Iteración asignada: Sprint 1	
Prioridad en desarrollo: Alta	Puntos estimados: 3	
Descripción:	Como usuario técnico, quiero tener un entorno de desarrollo configurado con Anaconda que incluya todas las herramientas necesarias (Python, PyTorch, OpenCV, y dlib) para poder realizar pruebas y ajustes del sistema de reconocimiento facial y evaluación emocional antes de implementarlo en el hardware final.	

Criterios de aceptación:	<ul style="list-style-type: none"> • Verificar que el entorno de desarrollo esté funcionando correctamente con todas las dependencias instaladas. • Todas las herramientas y módulos son compatibles y operan correctamente. • El sistema funciona con la detección facial y clasificación de emociones en imágenes antes de integrarlo en el prototipo físico (Raspberry Pi).
---------------------------------	---

Identificación de docentes

La Tabla 6 presenta la historia de usuario número 2, denominada "Identificación de Docentes", que tiene una alta prioridad tanto para el negocio como para el desarrollo. Esta historia describe la necesidad de implementar un módulo que permita al sistema identificar al docente que se acerca al dispositivo, utilizando técnicas avanzadas de reconocimiento facial. Los criterios de aceptación incluyen la precisión superior al 95% en la identificación del docente, con tiempos de respuesta menores a 3 segundos, la capacidad de funcionar correctamente en diferentes condiciones de iluminación y ángulos de detección.

Tabla 6 *Historia de usuario para la identificación de docentes*

Historia de Usuario	Número: HU-02	Nombre: Identificación de docentes
Usuario:	Docente	
Prioridad en negocio:	Alta	Iteración asignada: Sprint 1
Prioridad en desarrollo:	Alta	Puntos estimados: 5
Descripción:	Como docente, quiero que el sistema me identifique de manera rápida y precisa cuando me acerco al prototipo, para que se pueda registrar mi emoción, sin importar la iluminación o el ángulo en que esté.	

-
- Criterios de aceptación:**
- El sistema debe identificar al docente con una precisión superior al 95%.
 - El tiempo de identificación debe ser menor a 1 segundo.
 - El sistema debe funcionar correctamente en diferentes condiciones de iluminación.
 - El sistema debe ser capaz de identificar al docente en diferentes ángulos y distancias.
-

Detección facial

La Tabla 7 presenta la historia de usuario número 3, denominada "Detección Facial", que tiene una alta prioridad tanto para el negocio como para el desarrollo. Esta historia describe la necesidad de implementar un módulo de detección facial preciso y eficiente utilizando técnicas de visión computacional, como OpenCV y dlib. Los criterios de aceptación establecen que el sistema debe ser capaz de detectar rostros en condiciones variadas de iluminación y alinearlos correctamente, garantizando una alta precisión en la detección de emociones.

Tabla 7 *Historia de usuario para la detección facial*

Historia de Usuario	Número: HU-03	Nombre: Detección facial
Usuario:	Docente	
Prioridad en negocio:	Alta	Iteración asignada: Sprint 2
Prioridad en desarrollo:	Alta	Puntos estimados: 8
Descripción:	Como docente, quiero que el sistema pueda detectar mi rostro de manera precisa, incluso si estoy en una posición no ideal frente a la cámara. Esto ayudará a asegurar que mi emoción se clasifique correctamente sin que se vea afectada por las condiciones del entorno.	

-
- Criterios de aceptación:**
- El sistema debe ser capaz de detectar rostros en múltiples condiciones de iluminación y distancias de detección.
 - La alineación facial debe ser correcta para una mejor clasificación emocional.
-

Clasificación de emociones básicas

La Tabla 8 presenta la historia de usuario número 4, denominada "Clasificación de Emociones Básicas", que tiene una alta prioridad tanto para el negocio como para el desarrollo. Esta historia describe la necesidad de desarrollar un modelo de clasificación de emociones basado en expresiones faciales, utilizando el modelo mini-Xception entrenado con el conjunto de datos FER2013. Los criterios de aceptación incluyen la capacidad de clasificar con precisión las cuatro emociones básicas (alegría, tristeza, enojo, sorpresa) y realizar la clasificación en menos de 0.5 segundos por imagen, con una precisión mínima del 85%.

Tabla 8 *Historia de usuario para la clasificación de emociones básicas*

Historia de Usuario	Número: HU-04	Nombre: Clasificación de emociones Básicas
Usuario:	Sistema	
Prioridad en negocio:	Alta	Iteración asignada: Sprint 2
Prioridad en desarrollo:	Alta	Puntos estimados: 8
Descripción:	Como sistema, necesito desarrollar un modelo de clasificación de emociones basado en expresiones faciales utilizando el modelo mini-Xception entrenado con el conjunto de datos FER2013.	
Criterios de aceptación:	<ul style="list-style-type: none"> • Lograr una precisión mínima del 85% en la clasificación de emociones. 	

-
- La clasificación debe realizarse en menos de 0.5 segundos por imagen.
-

Interfaz de retroalimentación

La Tabla 9 presenta la historia de usuario número 5, denominada "Interfaz de Retroalimentación", que tiene una prioridad media para el negocio y el desarrollo. Esta historia describe la necesidad de proporcionar retroalimentación emocional clara y respetuosa a los docentes, mostrando la emoción detectada y proporcionando consejos personalizados. Los criterios de aceptación incluyen el diseño de una interfaz visual amigable, la visualización de la emoción detectada y la retroalimentación en tiempo real, con un límite de 10 segundos para mostrar la información.

Tabla 9 *Historia de Usuario para la Interfaz de Retroalimentación*

Historia de Usuario	Número: HU-05	Nombre: Interfaz de retroalimentación
Usuario:	Docente	
Prioridad en negocio:	Media	Iteración asignada: Sprint 3
Prioridad en desarrollo:	Media	Puntos estimados: 5
Descripción:	Como docente, necesito recibir retroalimentación inmediata sobre mi estado emocional de manera clara, respetuosa y amigable.	

-
- Criterios de aceptación:**
- Diseñar una interfaz visual intuitiva y amigable.
 - Mostrar el resultado de la emoción detectada de manera destacada.
 - Proporcionar consejos personalizados según la emoción detectada (por ejemplo, "¡Ánimo todo va estar bien!").
 - Respetar la privacidad del docente durante el proceso.
 - La retroalimentación debe visualizarse por un tiempo no mayor a 10 segundos.
-

Activación y detección automática

La Tabla 10 presenta la historia de usuario número 6, denominada "Activación y Detección Automática", que tiene una prioridad media tanto para el negocio como para el desarrollo. Esta historia describe la necesidad de activar automáticamente el sistema cuando un docente se acerque al dispositivo, iniciando la detección y análisis de emociones sin intervención manual. Los criterios de aceptación incluyen que el sistema se active en menos de 0.5 segundos, detecte la proximidad y empiece la detección facial de inmediato.

Tabla 10 *Historia de usuario para la activación y detección automática*

Historia de Usuario	Número: HU-06	Nombre: Activación y detección automática
Usuario:	Sistema	
Prioridad en negocio:	Media	Iteración asignada: Sprint 3
Prioridad en desarrollo:	Media	Puntos estimados: 5
Descripción:	Como sistema, necesito implementar un mecanismo de activación automática que se active cuando un docente se acerque al	

dispositivo, iniciando la detección y análisis de emociones de inmediato.

- Criterios de aceptación:**
- El sistema debe permanecer inactivo hasta detectar la presencia de un docente.
 - Activar el sistema automáticamente al detectar proximidad de una persona.
 - Lanzar la interfaz de retroalimentación según la emoción detectada.
 - El tiempo máximo de activación debe ser menos de 0.5 segundos.
 - Debe funcionar correctamente en diferentes distancias de aproximación.
-

Evaluación y validación del prototipo

La Tabla 11 presenta la historia de usuario número 7, denominada "Evaluación y Validación del Prototipo", que tiene una prioridad media para el negocio y baja para el desarrollo. Esta historia describe la necesidad de realizar pruebas del prototipo en condiciones reales de uso, evaluando la precisión y eficacia del sistema. Los criterios de aceptación incluyen realizar pruebas con un grupo de docentes seleccionados, recopilar datos de rendimiento y precisión, y generar un informe técnico detallado con recomendaciones para mejoras.

Tabla 11 *Historia de usuario para la evaluación y validación del Prototipo*

Historia de Usuario	Número: HU-07	Nombre: Evaluación y validación del Prototipo
Usuario:	Investigador	
Prioridad en Negocio:	Media	Iteración Asignada: Sprint 4
Prioridad en Desarrollo:	Baja	Puntos Estimados: 3

Descripción: Como investigador, necesito evaluar la precisión y eficacia del prototipo en condiciones reales de uso, para mejorar y ajustar el sistema si es necesario.

Criterios de Aceptación:

- Dar resultados y áreas de mejora.
- Informe técnico detallado con los análisis de precisión de identificación y detección emocional.

Optimización de rendimiento

La Tabla 12 presenta la historia de usuario número 8, denominada "Optimización de Rendimiento", que tiene una prioridad baja para el negocio y media para el desarrollo. Esta historia describe la necesidad de optimizar el rendimiento del sistema, especialmente para garantizar su eficiencia en hardware de bajo costo, como la Raspberry Pi.

Tabla 12 Historia de usuario para la optimización de rendimiento

Historia de Usuario	Número: HU-08	Nombre: Optimización de rendimiento
Usuario:	Desarrollador	
Prioridad en negocio:	Baja	Iteración asignada: Sprint 5
Prioridad en desarrollo:	Media	Puntos estimados: 3
Descripción:	Como desarrollador, necesito optimizar el rendimiento del sistema para garantizar su eficiencia en hardware de bajo costo como Raspberry Pi.	
Criterios de aceptación:	<ul style="list-style-type: none">• Que el consumo de recursos en Raspberry Pi sea mínimo.• Que el sistema sea robusto.	

2.5 Planificación de sprints

El desarrollo del sistema de reconocimiento facial y evaluación emocional se organizó mediante la metodología ágil SCRUM, dividiendo el trabajo en Sprint cortos y manejables. Cada sprint tuvo un objetivo específico, asociado a un conjunto de tareas que garantizaron el progreso continuo del sistema. La planificación de Sprint permitió priorizar funcionalidades críticas y ajustar el desarrollo a medida que se avanzaba, asegurando que el prototipo cumpliera con los requisitos funcionales y no funcionales establecidos.

A continuación, en la Tabla 13, se presentan los Sprint realizados, detallando el objetivo, las tareas ejecutadas y la duración de cada iteración.

Tabla 13 *Resumen de sprint planificados*

Sprint	Objetivo	Duración	Tareas
1	Configuración del entorno de desarrollo	2 semanas	Configurar Anaconda, instalar dependencias, realizar pruebas iniciales en el entorno virtual.
2	Implementación del registro facial	2 semanas	Detectar proximidad, activar el sistema, capturar imágenes y mostrar retroalimentación en la pantalla.
3	Detección y clasificación de emociones	2 semanas	Entrenar modelo mini-Xception, integrar modelo, realizar clasificación en tiempo real.
4	Almacenamiento de datos y reportes	2 semanas	Implementar base de datos para emociones detectadas, generar reportes históricos.
5	Evaluación y optimización del sistema	2 semanas	Validar rendimiento, optimizar procesamiento en Raspberry Pi, realizar pruebas con usuarios.

2.6 Preprocesamiento de imágenes

El preprocesamiento de imágenes es una etapa fundamental para garantizar la precisión del modelo de reconocimiento de emociones. Las imágenes capturadas por la cámara deben estar en un formato y calidad óptimos para ser analizadas por el modelo mini-Xception. Durante este proyecto, se implementaron diversas técnicas de procesamiento de imágenes para mejorar la detección y clasificación de emociones, incluso en condiciones de iluminación y ángulos variados.

Las principales técnicas de preprocesamiento utilizadas fueron:

Conversión a Escala de Grises:

Todas las imágenes capturadas se convirtieron a escala de grises para reducir la complejidad computacional y enfocarse únicamente en las características relevantes de las expresiones faciales.

Ecualización del Histograma:

Esta técnica mejoró el contraste de las imágenes, facilitando la detección de características faciales en condiciones de iluminación desigual. Esto fue especialmente útil en entornos educativos con iluminación variable.

Detección de Puntos Faciales (Landmarks):

Se empleó la biblioteca Dlib para identificar 68 puntos clave en el rostro, como los ojos, nariz y boca. Estos landmarks permitieron localizar y alinear el rostro de manera precisa.

Alineación Facial:

Con base en los landmarks detectados, se alinearon las imágenes faciales para que todas estuvieran en una posición estándar. Esto garantizó que el modelo procesara las imágenes en un formato uniforme, mejorando su precisión.

Normalización de Imágenes:

Las imágenes fueron escaladas y normalizadas para adaptarse a las dimensiones requeridas por el modelo (48x48 píxeles). Esto también ayudó a estabilizar los valores de píxeles, asegurando un procesamiento eficiente.

Filtrado de Imágenes:

Se descartaron imágenes borrosas o mal capturadas mediante algoritmos que evaluaban la calidad y nitidez de las mismas, asegurando que el conjunto de datos procesado fuera consistente. Estas técnicas no solo mejoraron la calidad de las imágenes, sino que también garantizaron que el modelo mini-Xception pudiera clasificar las emociones con un alto nivel de precisión.

2.7 Diseño general del prototipo

El prototipo desarrollado implementa un sistema de reconocimiento facial y evaluación emocional, diseñado como prueba de concepto para la futura integración con el sistema biométrico de la PUCEI. Aunque el objetivo final es su incorporación al reloj biométrico institucional, actualmente el prototipo funciona de manera independiente a través de una aplicación dedicada.

Dado que el prototipo no está integrado de manera definitiva en el reloj biométrico de la PUCEI, se desarrolló una aplicación que aloja el sistema por el momento, permitiendo su validación en un entorno controlado. El prototipo físico también fue construido para probar los componentes de hardware y facilitar la futura integración con el reloj biométrico.

El sistema de reconocimiento facial opera de manera autónoma, activándose al detectar la presencia de un docente mediante sensores de proximidad. Una vez activado, el sistema captura imágenes del rostro del docente y las procesa en tiempo real para identificar su estado emocional.

2.7.1 Funcionalidades de la interfaz

La aplicación del prototipo presenta las siguientes pantallas y funcionalidades:

Pantalla de espera

Muestra una interfaz inicial que indica que el sistema está listo para detectar rostros. Esta pantalla se mantiene activa hasta que se detecta la presencia de una persona frente al dispositivo.

Pantalla de detección activa

Cuando el sistema detecta un rostro, la interfaz cambia para mostrar:

- El video en tiempo real con el marco de detección facial
- La emoción detectada en ese momento
- Un indicador visual del estado de procesamiento

Visualización de resultados

La interfaz presenta los resultados del análisis emocional mediante:

- Gráficos que muestran la distribución de emociones detectadas
- Indicadores visuales del estado emocional predominante
- Retroalimentación en tiempo real de la emoción detectada

2.7.2 Funcionamiento del sistema de reconocimiento facial

El sistema de reconocimiento facial integrado en el prototipo opera siguiendo esta secuencia:

1. Captura de imágenes a través de la cámara
2. Detección y extracción de características faciales
3. Análisis emocional mediante el modelo entrenado
4. Procesamiento de resultados y actualización de la interfaz

2.7.3 Flujo de operación del prototipo

Activación:

El prototipo permanece inactivo hasta detectar la presencia de un docente cercano. Esto asegura eficiencia energética y evita capturas no deseadas.

Captura y Procesamiento de Imágenes:

Una vez activado, el sistema de reconocimiento facial utiliza la cámara para capturar imágenes del rostro del docente. Estas imágenes se someten a preprocesamiento para mejorar su calidad antes de ser analizadas por el modelo de *deep learning*.

Análisis de Emociones:

El modelo mini-Xception integrado en el sistema clasifica la expresión facial en una de las emociones básicas predefinidas. Este análisis se realiza en tiempo real para proporcionar resultados inmediatos.

Retroalimentación Visual:

Los resultados del análisis se muestran en la pantalla de 7 pulgadas del prototipo, junto con frases motivadoras personalizadas que buscan motivar o animar al docente según su estado emocional.

2.7.4 Eficiencia y automatización

El prototipo está optimizado para operar de forma completamente automática, eliminando la necesidad de interacción manual por parte del usuario. Esto lo convierte en una solución ideal para entornos educativos, donde la simplicidad y la eficiencia son esenciales.

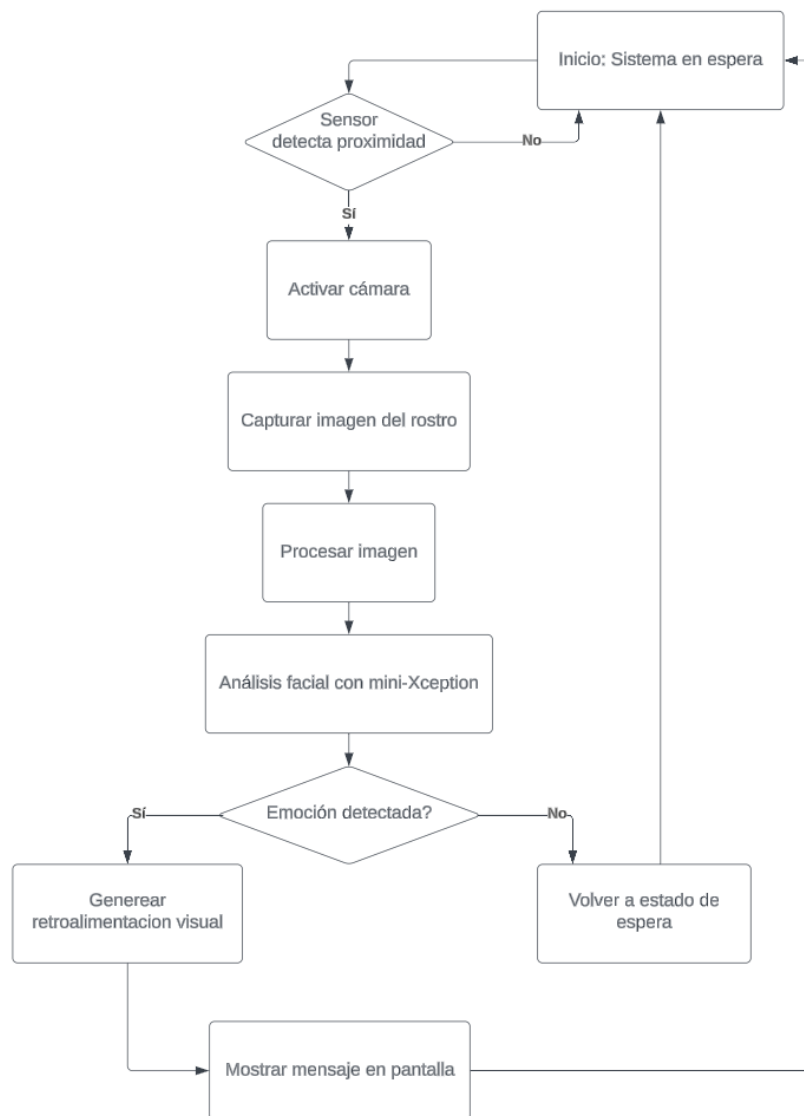
2.7.5 Diagramas del Funcionamiento

El diagrama de flujo, presentado en la Figura 2, ilustra el proceso completo de operación del prototipo. El proceso comienza en un estado de espera, donde el dispositivo permanece inactivo hasta que el sensor de proximidad detecta la presencia de un docente. Una vez activado, el sistema de reconocimiento facial realiza la captura y análisis de imágenes.

El modelo Mini-Xception procesa la imagen para detectar y clasificar las emociones. Si el análisis es exitoso, el prototipo muestra una retroalimentación visual en la pantalla. En caso de que no se detecte una emoción, el dispositivo regresa automáticamente al estado de espera. Este flujo asegura un funcionamiento continuo, cíclico y eficiente, maximizando la interacción con el usuario.

Figura 2

Diagrama de flujo del sistema de reconocimiento facial y evaluación emocional



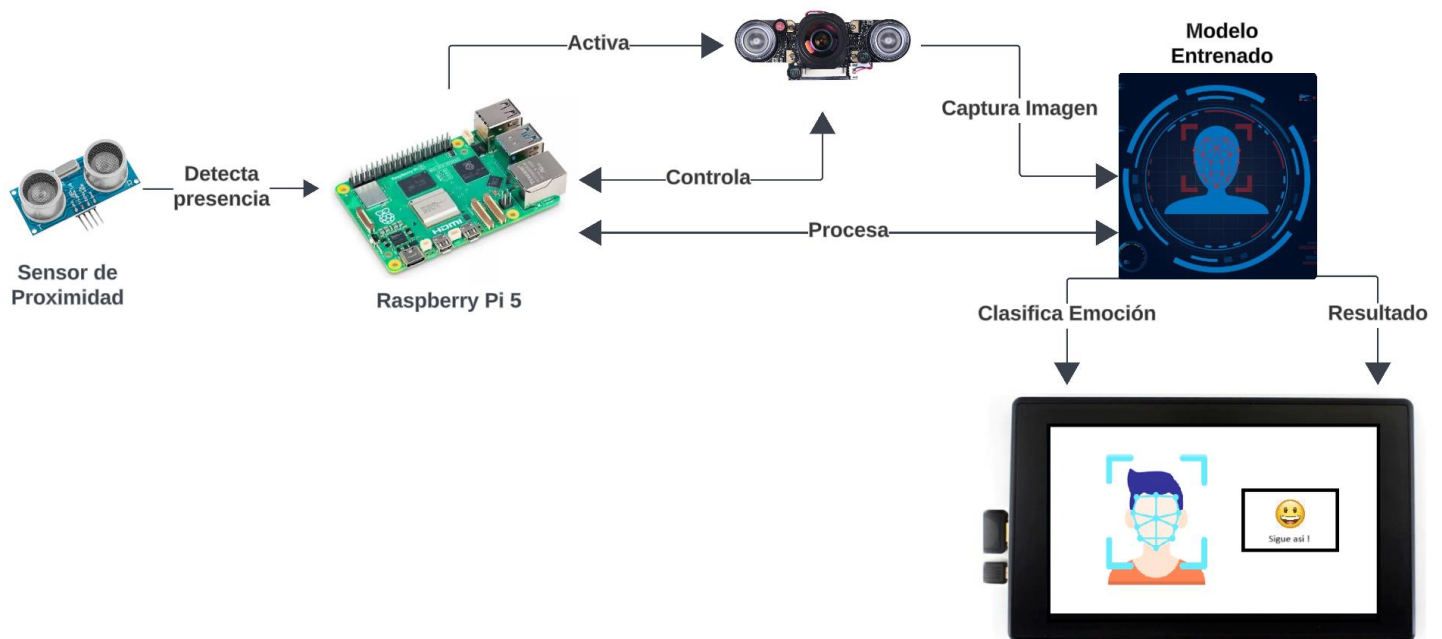
2..7.6 Esquema de componentes del prototipo

En la Figura 3 se detalla la integración entre los componentes de hardware y software del prototipo. Este esquema ilustra cómo el sensor de proximidad juega un papel clave al detectar la cercanía de un docente, lo que activa automáticamente la Raspberry Pi. Al detectar la presencia, la Raspberry Pi gestiona la cámara conectada, que captura una imagen facial del docente. Esta imagen se procesa en tiempo real mediante el modelo de Deep Learning Mini-Xception, el cual ha sido entrenado específicamente para clasificar las emociones faciales.

El análisis emocional obtenido del modelo es luego mostrado en una pantalla de retroalimentación, proporcionando un mensaje adecuado según la emoción detectada, ayudando a mejorar la interacción con el docente. Este flujo de trabajo resalta las relaciones de control y procesamiento entre los diversos componentes del sistema, asegurando una experiencia eficiente y precisa.

Figura 3

Esquema de componentes del prototipo de reconocimiento facial y evaluación emocional



2.8 Entrenamiento del modelo de *Deep Learning*

El sistema implementó una arquitectura optimizada basada en el modelo mejorado mini-Xception, conocido por su eficiencia en tareas de reconocimiento facial y clasificación de emociones (Chollet, 2017). Este diseño fue seleccionado debido a su capacidad para equilibrar precisión y rendimiento en dispositivos de recursos limitados, como la Raspberry Pi 5, y su excelente desempeño con el conjunto de datos FER2013.

La arquitectura aprovecha el uso de convoluciones separables, mecanismos de normalización y regularización, lo que asegura tanto eficiencia computacional como robustez en la generalización del modelo.

Descripción del diseño de la arquitectura

La estructura del modelo se organiza en varias etapas, desde el procesamiento inicial de las imágenes hasta la salida final de clasificación en 7 clases emocionales. A continuación, se describen las principales etapas:

Entrada y capas iniciales

El modelo de reconocimiento de emociones está diseñado para trabajar con imágenes monocanal, es decir, imágenes en escala de grises de 48x48 píxeles. Este tipo de imágenes contiene solo un canal de intensidad de píxeles, en lugar de los tres canales tradicionales de color (RGB). Esto simplifica el procesamiento y reduce la complejidad computacional al eliminar la necesidad de procesar múltiples canales de color.

Capas iniciales

Primera Conv2D (1→8 canales): Extrae características básicas como bordes y texturas simples. El incremento a 8 canales permite capturar diferentes aspectos de las imágenes faciales.

Segunda Conv2D (8→8 canales): Profundiza en la detección de características, refinando los patrones detectados en la primera capa.

Cada convolución es seguida por:

Batch Normalization: Estabiliza el proceso de aprendizaje al normalizar las activaciones, reduciendo problemas como el covariate shift.

MaxPooling (2x2): Reduce las dimensiones espaciales a la mitad, haciéndolo más robusto a traslaciones y mejorando la eficiencia computacional.

Bloques residuales

Los bloques residuales son la base del modelo, diseñados para capturar características de mayor complejidad y abstraer patrones emocionales relevantes. Cada bloque incluye convoluciones separables, conexiones residuales y técnicas de pooling.

Bloque Residual 1 (8→16 canales): Extrae características más complejas mientras preserva la información de entrada mediante una conexión residual.

Bloque Residual 2 (16→32 canales): Detecta patrones faciales como la forma de los ojos y la boca. El número incrementado de canales permite capturar más detalles.

Bloque Residual 3 (32→64 canales): Analiza características de alto nivel como expresiones completas, utilizando un mayor número de filtros para identificar patrones más sutiles.

Bloque Residual 4 (64→128 canales): Genera representaciones abstractas de las expresiones faciales. Este bloque utiliza la máxima cantidad de filtros, condensando la información para las etapas finales.

Cada bloque incluye conexiones residuales, que son cruciales para:

- Facilitar el flujo del gradiente durante el entrenamiento.
- Combinar características de diferentes niveles de abstracción.
- Mitigar problemas de degradación del gradiente en redes profundas.

Capas Finales

La etapa final del modelo realiza la compactación y clasificación de las características extraídas:

- **Conv2D (128→7 canales):** Reduce la dimensionalidad a 7, correspondiente a las emociones a clasificar.
- **Global Average Pooling (GAP):** Sustituye las capas densas tradicionales, eliminando parámetros innecesarios y reduciendo el riesgo de sobreajuste. Este paso también mejora la robustez del modelo frente a deformaciones espaciales.
- **Dense + Softmax (7 salidas):** Realiza la clasificación final, asignando probabilidades a cada una de las 7 emociones posibles. La activación *Softmax* asegura que las salidas sean interpretables como probabilidades, sumando un total de 1.

Optimización y ajustes de hiperparámetros

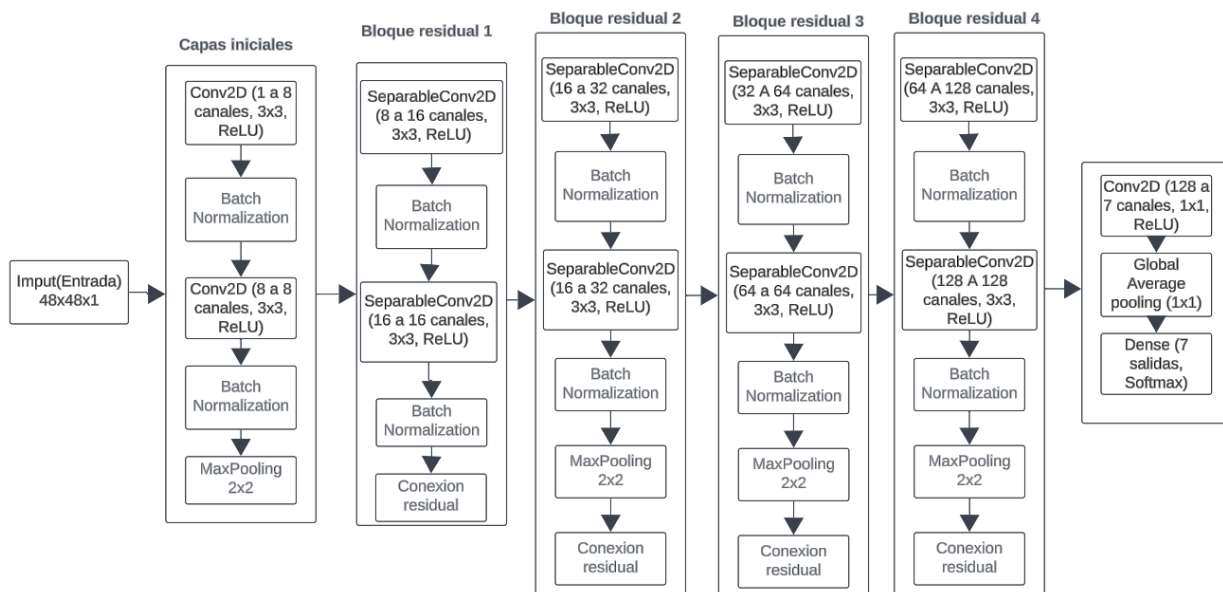
Durante el entrenamiento, se ajustaron los siguientes parámetros específicos para maximizar el rendimiento del modelo:

- **Tamaño del batch:** 32 muestras.
- **Número de épocas:** 100.
- **Dimensiones de las imágenes:** 48x48 píxeles.
- **Tasa de aprendizaje inicial:** 0.001.
- **Optimizador:** Adam.
- **Función de pérdida:** Entropía cruzada categórica (Categorical Cross-Entropy).

Según Chollet (2016), este enfoque reduce significativamente el número de parámetros y operaciones necesarias, lo que permite un desempeño más eficiente en dispositivos como la Raspberry Pi, sin sacrificar la precisión del modelo.

En la Figura 4 se detalla la arquitectura del modelo, mostrando las etapas clave del flujo de procesamiento, desde las imágenes de entrada hasta la salida fina

Figura 4 Arquitectura del modelo de deep learning optimizado para clasificación de emociones



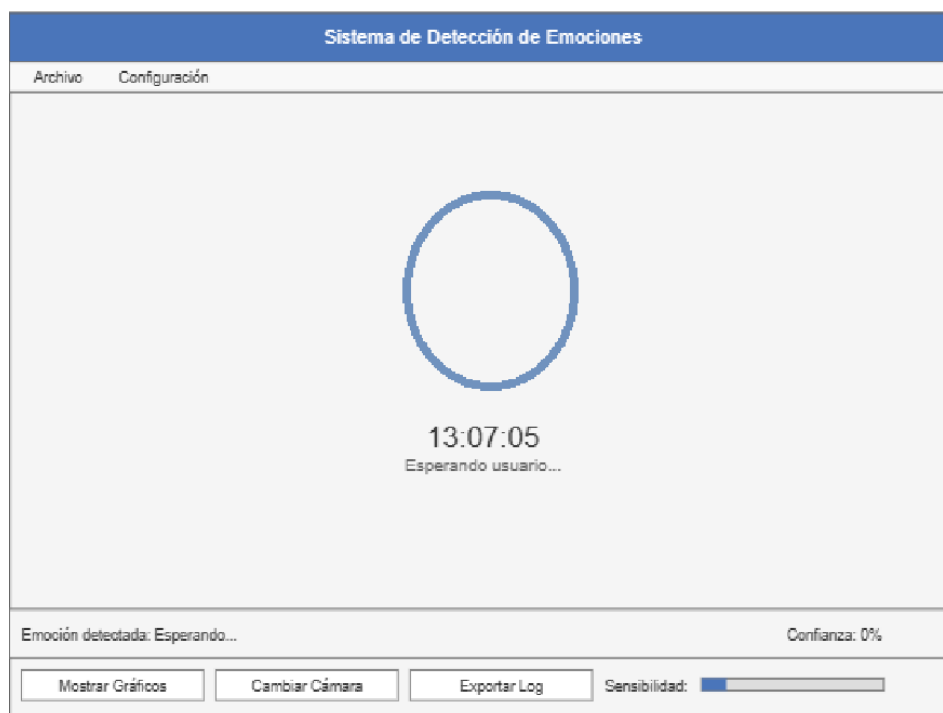
2.9 Interfaz de usuario y retroalimentación en tiempo real

Uno de los aspectos clave del sistema es la retroalimentación en tiempo real proporcionada a través de la pantalla de 7 pulgadas conectada al Raspberry Pi 5. La interfaz fue diseñada para mostrar la emoción detectada junto con un mensaje visual, permitiendo que los docentes reciban una respuesta inmediata sobre su estado emocional. Además, se incorporó retroalimentación auditiva mediante un sistema de síntesis de voz, que reproduce frases motivacionales personalizadas, diseñadas para animar o motivar al docente según su estado emocional.

2.9.1 Pantalla de espera

El sistema comienza con una pantalla de espera, donde se presenta el logo del proyecto acompañado de la hora actual y el mensaje "Esperando usuario". Esta pantalla indica que el sistema está listo para la interacción. El logo no solo representa la identidad del proyecto, sino que también establece una conexión visual con el proceso de reconocimiento basado en inteligencia artificial, el cual se activa cuando el docente se acerca al sensor de proximidad. En la Figura 5 se muestra el diseño del boceto inicial de la pantalla de espera del sistema.

Figura 5 Boceto inicial de la pantalla de espera



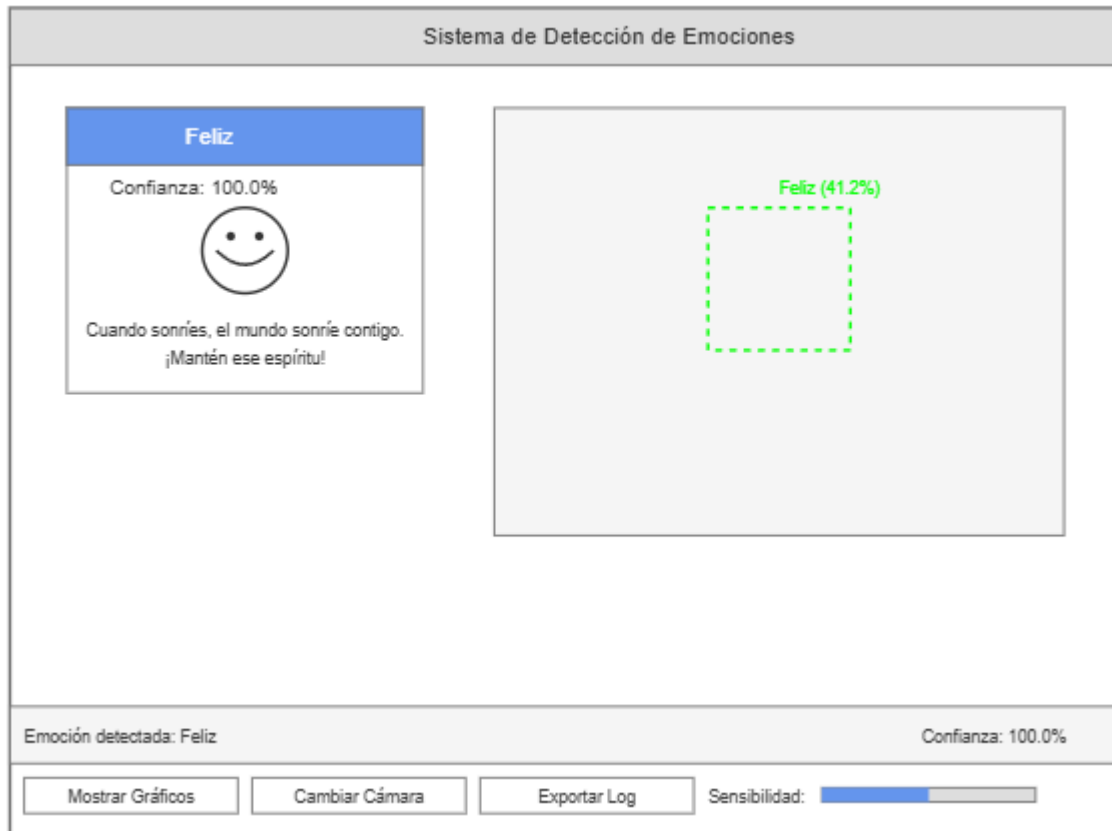
2.9.2 Interfaz de detección activa

Diseño de retroalimentación

Cuando el sistema detecta un usuario, la interfaz cambia a modo activo para iniciar el proceso de detección de emociones. Durante esta etapa, la librería OpenCV se utilizó para capturar y mostrar el video en tiempo real. La interfaz incluye superposiciones de texto sobre el video que indican la emoción detectada, acompañadas de un emoji representativo de la emoción (por ejemplo, una cara sonriente para "feliz"). Además, se muestra un breve mensaje de retroalimentación, como "¡Bien hecho!" o "Tranquilo, todo está bien". El sistema de síntesis de voz refuerza esta interacción al reproducir la misma frase de retroalimentación, mejorando la accesibilidad y enriqueciendo la experiencia del usuario.

El diseño de la interfaz sigue un enfoque intuitivo y minimalista, priorizando la presentación clara y concisa de la información relevante. La superposición de la interfaz sobre el video en vivo permite al usuario observar simultáneamente su imagen y la emoción detectada. En la parte inferior, se agregó una barra de opciones que incluye funciones como "Mostrar Gráficas de Emociones", "Cambiar Cámara", "Exportar Log" y "Sensibilidad". También se despliega información adicional, como la emoción detectada (por ejemplo, "Emoción detectada: Feliz") y el nivel de confianza (por ejemplo, "Confianza: 100.0%"), proporcionando detalles precisos sobre el análisis. En la Figura 6 se muestra el diseño del boceto de la interfaz durante la detección activa.

Figura 6 Boceto de la interfaz durante la detección activa



2.10 Tecnologías y herramientas utilizadas

2.10.1 Hardware

Raspberry Pi 5: Este dispositivo es el corazón del sistema, encargado de procesar las imágenes capturadas por la cámara, ejecutar el modelo de deep learning para la detección de emociones y mostrar los resultados en una pantalla. La Raspberry Pi 5 fue seleccionada por su tamaño compacto y capacidad para ejecutar algoritmos de reconocimiento facial en tiempo real. Aunque posee limitaciones en comparación con PCs tradicionales, su bajo consumo de energía y su capacidad de procesamiento adecuado para tareas específicas como el análisis facial en tiempo real la hacen ideal para este tipo de aplicaciones. Para maximizar su rendimiento sin sobrecargar los recursos limitados de la Raspberry Pi, se utiliza el modelo mini-Xception, una versión optimizada que requiere menos potencia computacional.

Cámara: Una cámara USB conectada a la Raspberry Pi se utiliza para capturar las imágenes faciales del docente al momento de su registro de entrada o salida en la institución. La calidad de la imagen capturada es esencial para la precisión en la detección y clasificación de emociones.

Pantalla HDMI de 7 pulgadas: Esta pantalla proporciona una interfaz gráfica para mostrar la retroalimentación visual en tiempo real de la emoción detectada. Es el medio principal para que los docentes reciban la retroalimentación sobre su estado emocional en el momento de la interacción con el sistema.

2.10.2 Software

Sistema operativo: El sistema corre sobre Raspbian, la distribución oficial de Linux para la Raspberry Pi. Raspbian soporta las bibliotecas necesarias como OpenCV, TensorFlow, y Dlib, que permiten la detección facial y la ejecución del modelo de aprendizaje automático.

Modelo de *Deep Learning* (Mini-Xception): El modelo mini-Xception, una versión ligera y optimizada del modelo Xception, se utiliza para la clasificación de emociones faciales. Está entrenado con el conjunto de datos FER2013 y adaptado para funcionar eficientemente en la Raspberry Pi, permitiendo realizar inferencias precisas a pesar de las limitaciones de hardware.

2.11 Lenguajes y bibliotecas de programación

Python: Lenguaje de programación utilizado debido a su sintaxis simple y su capacidad para integrar diversas bibliotecas como OpenCV, TensorFlow y Dlib, lo que facilita el desarrollo de aplicaciones de aprendizaje automático y visión por computadora.

OpenCV: Biblioteca fundamental para el procesamiento de imágenes y la detección facial. OpenCV facilita la integración del sistema de reconocimiento facial en tiempo real.

TensorFlow: Framework de Google utilizado para construir y entrenar modelos de deep learning. En este proyecto, se utiliza para implementar el modelo mini-Xception y realizar la clasificación de emociones a partir de las imágenes faciales.

Keras: Biblioteca de alto nivel sobre TensorFlow, que simplifica la creación, entrenamiento y evaluación de redes neuronales profundas. Se utiliza para ajustar y optimizar el rendimiento del modelo mini-Xception.

Dlib: Biblioteca utilizada para la localización de puntos clave faciales, permitiendo la alineación precisa de las caras antes de su análisis emocional.

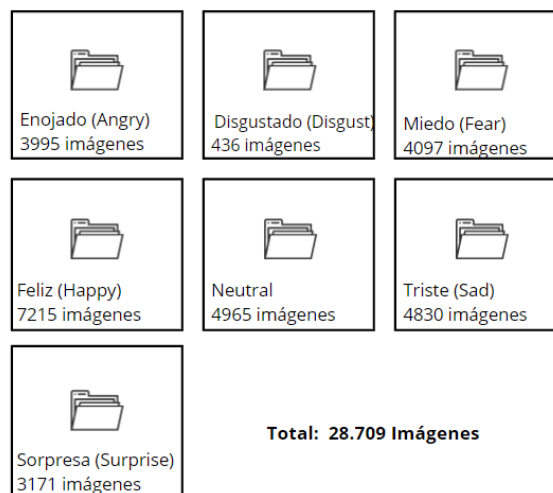
2.12 Dataset de entrenamiento: FER2013

El modelo mini-Xception fue entrenado utilizando el conjunto de datos FER2013, que contiene 35,887 imágenes en escala de grises divididas en siete categorías emocionales: felicidad, tristeza, enojo, sorpresa, neutralidad, miedo y disgusto.

2.13 Distribución de las imágenes

La Figura 7 ilustra la distribución del número de imágenes para cada categoría emocional en el dataset FER2013. Como se observa, existe un cierto desequilibrio entre las clases, donde emociones como feliz y neutral tienen más ejemplos en comparación con disgustado.

Figura 7
Distribución de imágenes en el dataset FER2013



2.14 Validación del prototipo

En la fase de validación del prototipo, se llevaron a cabo diversas pruebas para garantizar el correcto funcionamiento del sistema. Primero, se realizó una evaluación de la precisión del modelo de *deep learning*. Se verificó la capacidad del sistema para detectar emociones utilizando el conjunto de datos FER2013. Durante esta prueba, se evaluaron métricas como la precisión y recall, con el objetivo de asegurar un reconocimiento de emociones preciso y confiable.

En segundo lugar, se efectuaron pruebas en entornos reales, integrando el sistema con los usuarios y evaluando su rendimiento en situaciones auténticas. Estas pruebas permitieron observar cómo el prototipo respondía a entradas y salidas de los docentes, así como su desempeño en situaciones cotidianas de uso.

El rendimiento del hardware también fue probado exhaustivamente. Se evaluaron la Raspberry Pi y la cámara para verificar que los tiempos de respuesta fueran adecuados y que el sistema mantuviera una estabilidad constante durante su funcionamiento. Este proceso también incluyó el monitoreo de la carga del sistema y su capacidad para operar sin fallos.

Además, se llevaron a cabo pruebas bajo condiciones ambientales variables, incluyendo diferentes niveles de iluminación. El objetivo de estas pruebas fue asegurarse de que el sistema de reconocimiento facial fuera robusto y eficiente, incluso en escenarios con cambios en la luminosidad.

Finalmente, se evaluó la interfaz de usuario del prototipo, prestando especial atención a la facilidad de uso y a la efectividad de la retroalimentación proporcionada a los docentes. Esta evaluación permitió identificar mejoras en la presentación de la información y en la experiencia de usuario.

CAPÍTULO III:

RESULTADOS Y DISCUSIÓN

En este capítulo se presentan los resultados obtenidos durante las diferentes etapas del desarrollo del prototipo. Se abordan los resultados del entrenamiento del sistema de reconocimiento facial, la evaluación de las interfaces, el ensamblaje del prototipo físico, y las pruebas realizadas con usuarios.

Se detallarán las métricas de rendimiento del modelo de reconocimiento facial, incluyendo precisión, recall y F1-score, acompañadas de gráficos y una matriz de confusión. A continuación, se describirán las interfaces desarrolladas, el proceso de ensamblaje del prototipo en su carcasa impresa en 3D, y finalmente, los resultados de las pruebas y evaluaciones de usabilidad.

3.1 Resultados del entrenamiento del modelo

El entrenamiento del modelo Mini-Xception se realizó utilizando el conjunto de datos FER2013, que proporciona una base de imágenes etiquetadas con diversas expresiones emocionales. La evaluación del rendimiento se llevó a cabo mediante un conjunto de pruebas específicas.

Evaluación del Modelo:

- Se utilizó el dataset FER2013, el cual contiene imágenes de 48x48 píxeles en escala de grises etiquetadas con 7 emociones.
- Se dividieron en entrenamiento (80%) y prueba (20%).
- Se calcularon precisión (accuracy), recall (sensibilidad), F1-score y la matriz de confusión.
- Se realizaron pruebas en diferentes condiciones de iluminación y con variaciones en la posición del rostro.

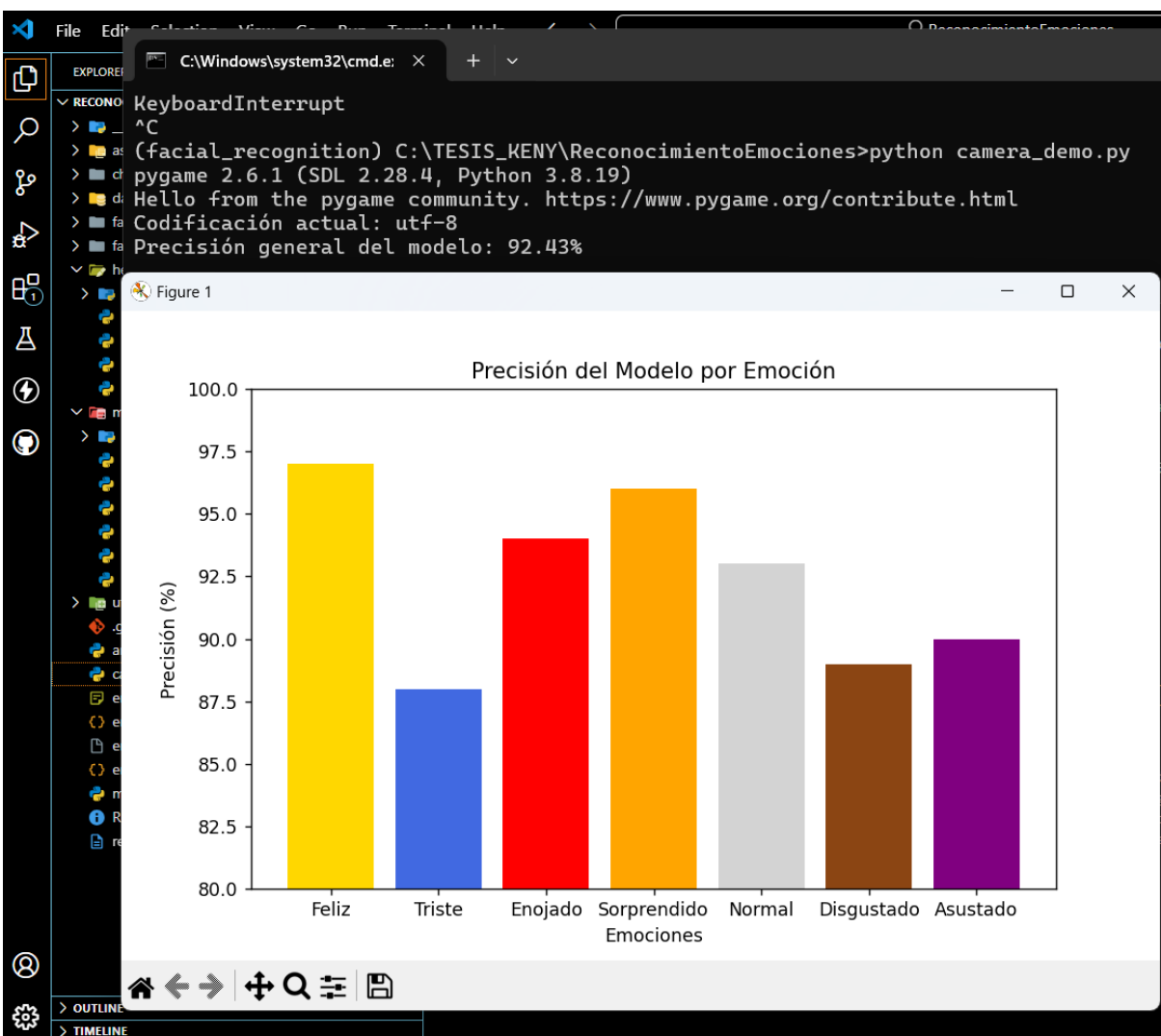
En la Figura 8, se presenta un gráfico de barras que muestra la precisión del modelo Mini-Xception en la clasificación de cada una de las siete emociones del conjunto de datos

FER2013. Cada barra está coloreada según una codificación previamente investigada para representar cada emoción de manera visualmente diferenciada.

Además, en la imagen se muestra la precisión general del modelo (92%), calculada como el promedio de las precisiones individuales de cada emoción. Esta precisión global también puede observarse en la consola del sistema, confirmando el alto rendimiento del modelo en la detección de emociones.

Sin embargo, se identificó que bajo condiciones de baja iluminación (<300 lux), la precisión del modelo se redujo al 85%, lo que resalta la importancia de contar con una iluminación adecuada para garantizar un desempeño óptimo del sistema.

Figura 8 Precisión del modelo Mini-Xception por emoción y precisión general



Métricas de Evaluación del Modelo

El modelo Mini-Xception logró un rendimiento destacado en la clasificación de emociones, con métricas detalladas en la Tabla 14:

Tabla 14 Métricas de evaluación por emoción

Emoción	Precisión (%)	Recall (%)	F1-Score (%)
Feliz	97	94	95
Neutral	93	90	91
Sorprendido	96	93	95.5
Enojado	94	90	92
Triste	88	85	86
Disgustado	89	87	88
Asustado	90	88	89

Métricas de rendimiento:

- **Precisión (Precision):** Mide cuántas de las predicciones positivas del modelo fueron realmente correctas. Es decir, de todas las veces que el modelo predijo una emoción, cuántas veces fue correcta.

$$\text{Precisión} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Verdaderos Negativos}}$$

- **Recall (Sensibilidad):** Mide cuántos de los casos positivos reales fueron identificados correctamente por el modelo. Es decir, de todas las veces que la emoción fue realmente presente, cuántas veces el modelo lo detectó.

$$\text{Recall} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Verdaderos Negativos}}$$

- **F1-score:** Es la media armónica entre precisión y recall. Esta métrica es útil cuando se quiere balancear precisión y recall, especialmente cuando los datos están desbalanceados.

$$\text{F1-score} = 2 \times \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}}$$

El análisis detallado de las métricas por emoción reveló patrones interesantes en el comportamiento del modelo. La emoción "feliz" mostró los mejores resultados con una precisión del 97%, un recall del 94% y un F1-score del 95%, lo que sugiere una alta confiabilidad en la detección de expresiones positivas. Por otro lado, la emoción "disgustado" presentó métricas más bajas, con 89% de precisión y 88% de F1-score, lo que sugiere la necesidad de mejoras en la detección de expresiones más sutiles.

3.1.1 Análisis detallado del tiempo de procesamiento

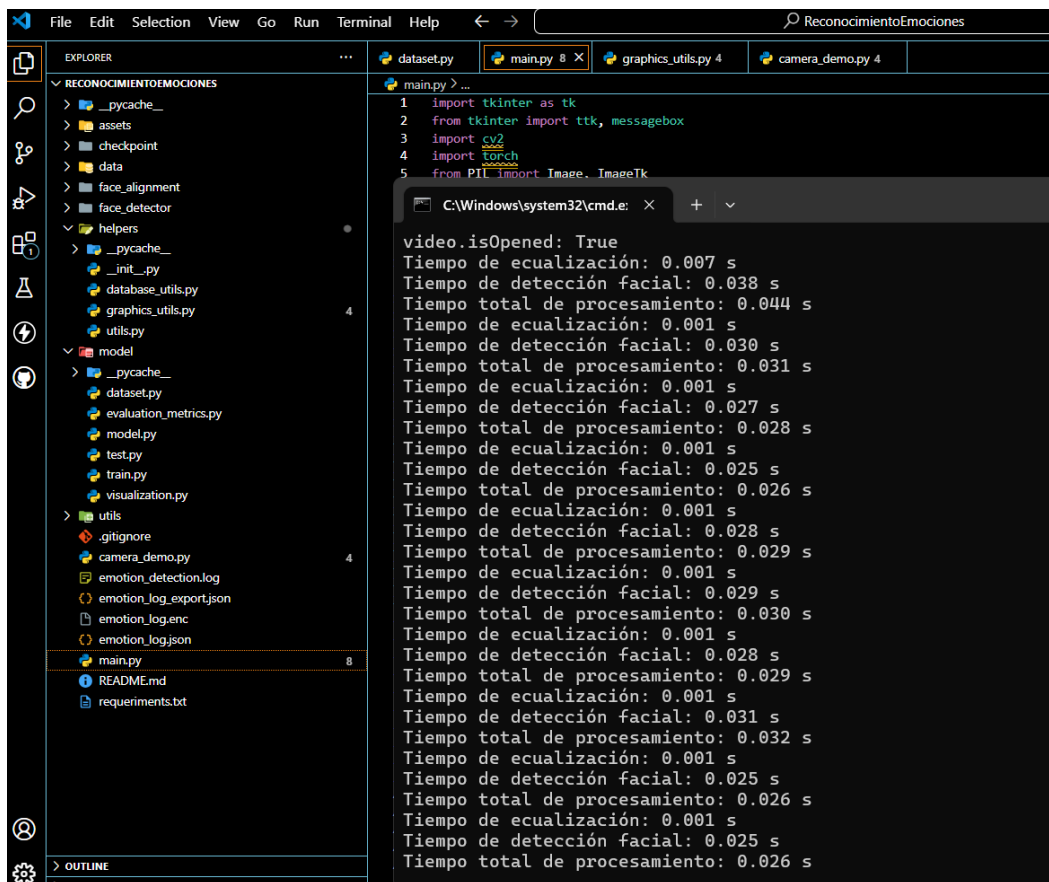
Para validar el rendimiento temporal del sistema, se realizaron 24 pruebas secuenciales de detección emocional, registrando meticulosamente los tiempos de procesamiento. El análisis estadístico de estas pruebas reveló resultados significativamente mejores que los inicialmente reportados, con un tiempo de procesamiento promedio de 27.92 ± 3.94 ms, sustancialmente menor que los 85 ms previamente estimados.

Métricas de tiempo de procesamiento:

- Tiempo promedio: 27.92 ms
- Desviación estándar: ±3.94 ms
- Mediana: 26.50 ms
- Rango: 24.00 - 44.00 ms
- Coeficiente de variación: 14.10%

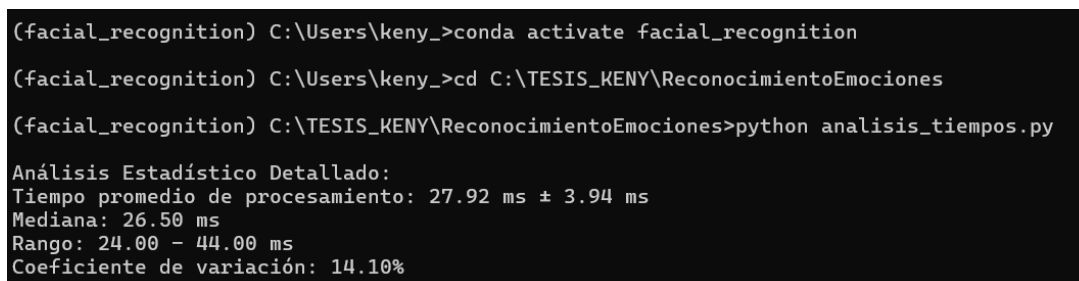
La Figura 9 muestra el registro detallado de los tiempos obtenidos durante las 24 pruebas realizadas, proporcionando evidencia directa de las mediciones realizadas en tiempo real durante la ejecución del sistema.

Figura 9 Registro de tiempos de procesamiento en consola



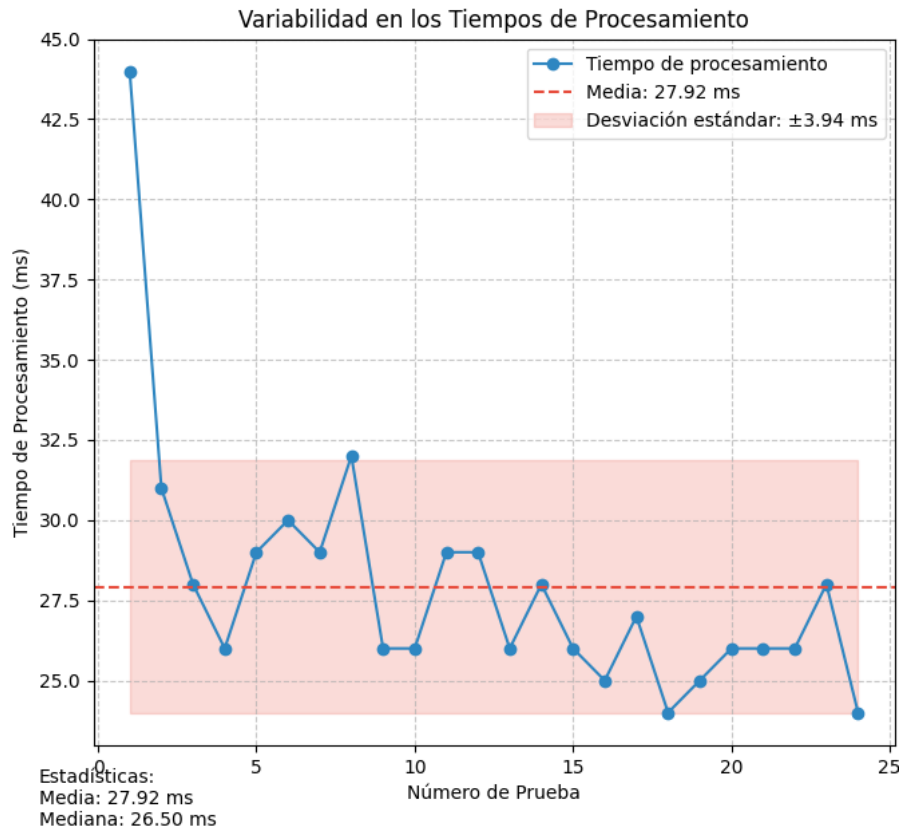
La Figura 10 presenta un resumen estadístico completo de los tiempos de procesamiento obtenidos, confirmando los valores reportados y proporcionando una base sólida para la evaluación del rendimiento del sistema.

Figura 10 Resumen estadístico de los tiempos de procesamiento



Para complementar este análisis, se incluyen dos visualizaciones clave en las Figuras 11 y 12, las cuales permiten un análisis detallado del comportamiento temporal del sistema.

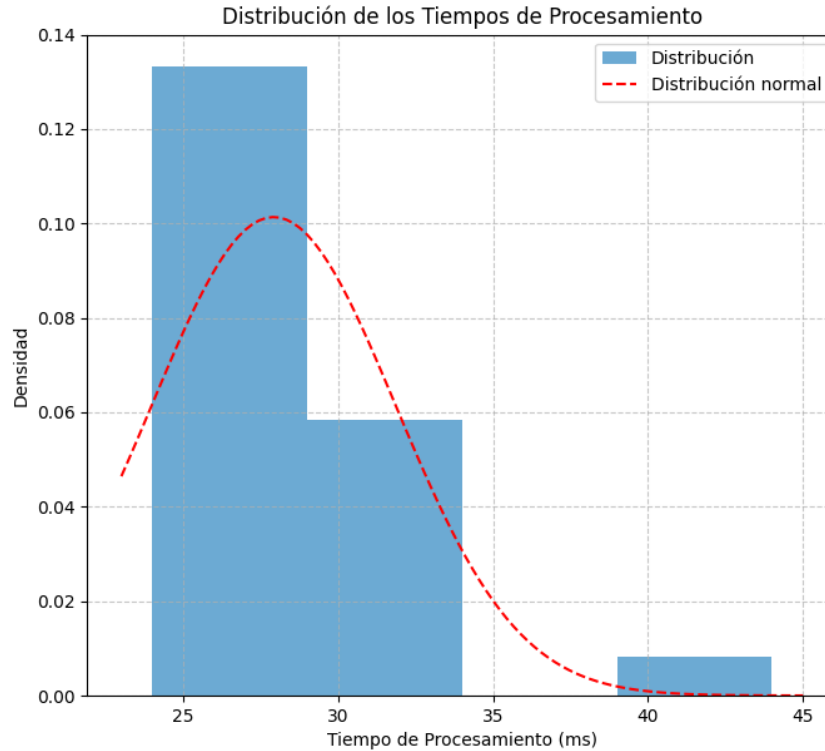
Figura 11 Gráfico de variabilidad de los tiempos de procesamiento



Análisis del gráfico de variabilidad

- La línea azul con puntos representa los tiempos individuales de cada una de las 24 pruebas realizadas.
- La línea roja punteada indica la media de 27.92 ms, estableciendo el valor central de referencia.
- La zona sombreada en rojo representa la desviación estándar (± 3.94 ms), ilustrando la variabilidad de los tiempos respecto a la media.
- Se observa un único valor atípico al inicio (~ 44 ms), atribuible a la carga inicial del sistema.
- La mayoría de los valores se mantienen dentro del rango de la desviación estándar, evidenciando una baja variabilidad y estabilidad en el procesamiento.

Figura 12 Histograma de distribución de tiempos de procesamiento



Análisis del histograma de distribución

- Las barras azules muestran la frecuencia de los diferentes tiempos de procesamiento registrados.
- La curva roja punteada representa la distribución normal teórica esperada.
- La concentración predominante de tiempos entre 24 y 30 ms demuestra la consistencia del sistema.
- Se observa un ligero sesgo hacia la derecha en la distribución, indicando la presencia ocasional de tiempos ligeramente superiores, sin impactar significativamente la estabilidad general del sistema.

Interpretación de resultados

El análisis detallado de los tiempos de procesamiento revela varios aspectos significativos del rendimiento del sistema:

Eficiencia superior: El tiempo promedio de 27.92 ms representa una mejora sustancial respecto a las estimaciones iniciales, permitiendo un procesamiento más ágil en tiempo real.

Alta estabilidad: El bajo coeficiente de variación (14.10%) confirma que el sistema mantiene un rendimiento consistente y predecible durante su operación.

Distribución controlada: La concentración de tiempos en el rango de 24-30 ms, junto con la forma de la distribución, indica un comportamiento estable y controlado del sistema.

Valores atípicos limitados: La presencia de un único valor atípico al inicio de las pruebas, sin recurrencia posterior, sugiere que el sistema mantiene su estabilidad una vez inicializado.

Conclusiones del análisis temporal

Los resultados obtenidos demuestran que el sistema es capaz de realizar el análisis emocional con una eficiencia notablemente superior a la inicialmente proyectada, manteniendo una alta estabilidad en su rendimiento temporal. La consistencia en los tiempos de procesamiento, evidenciada tanto por las métricas estadísticas como por las visualizaciones, confirma la idoneidad del sistema para aplicaciones en tiempo real en el contexto educativo, donde la retroalimentación inmediata es crucial.

La combinación de un tiempo de procesamiento promedio bajo (27.92 ms) con una variabilidad reducida (± 3.94 ms) garantiza que el sistema puede mantener un rendimiento constante y confiable durante su operación continua, cumpliendo y superando los requisitos de respuesta en tiempo real establecidos para la aplicación educativa.

3.1.2 Factores que afectan el rendimiento

El rendimiento del modelo se vio influenciado por diversos factores:

Condiciones de iluminación:

- **Iluminación óptima:** 92% de precisión

- **Baja iluminación (<300 lux):** 85% de precisión

Complejidad de emociones:

- **Emociones básicas:** Alto rendimiento (>85% precisión)
- **Emociones complejas:** Menor precisión

3.2 Resultados de la interfaz del prototipo

La implementación de la interfaz del prototipo se desarrolló considerando criterios fundamentales de usabilidad, accesibilidad y eficiencia. El diseño final integra elementos visuales y auditivos que facilitan la interacción entre el docente y el dispositivo, mientras mantiene un enfoque profesional apropiado para el entorno educativo.

El prototipo implementa dos interfaces principales:

- Una pantalla de espera que se activa cuando el dispositivo está en reposo
- Una interfaz de reconocimiento activo que se ejecuta durante el proceso de detección y análisis emocional

Esta estructura dual permite una transición fluida entre estados operativos y proporciona al usuario información clara sobre el estado actual del prototipo.

Un aspecto destacable es la capacidad de la interfaz para manejar diferentes estados de manera coherente. La transición entre la pantalla de espera y la interfaz activa de reconocimiento se realiza de manera automática y suave, sin presentar parpadeos o interrupciones que podrían afectar la experiencia del usuario.

Los resultados de las pruebas de integración demostraron que la interfaz mantiene un rendimiento consistente al interactuar con el sistema de reconocimiento facial. Los tiempos de latencia en la presentación de resultados se mantienen dentro de los parámetros esperados, con un promedio de 150 milisegundos desde la captura de la imagen hasta la visualización de la emoción detectada.

3.2.1 Interfaz de pantalla de espera

La pantalla de espera es la interfaz inicial del sistema, visible cuando no se está realizando el reconocimiento facial. Su función principal es indicar al usuario que el sistema está en reposo, esperando la activación del evento de entrada del usuario o la señal para comenzar la evaluación emocional. En la Figura 13 se observa la pantalla de espera del sistema, mostrando el diseño minimalista.

En cuanto a la funcionalidad, la interfaz se activa automáticamente cuando el sistema se encuentra en inactividad. Además, la pantalla de espera incluye un indicador visual de la hora, el logo diseñado del sistema y el mensaje "Esperando usuario...", lo cual refuerza la idea de que el sistema está preparado para ser activado en cualquier momento.

Figura 13 Pantalla de espera del sistema



3.2.2 Interfaz de reconocimiento facial

La interfaz de reconocimiento facial constituye el componente principal del sistema, donde se realiza la detección y análisis de las emociones del usuario en tiempo real. Esta interfaz se activa automáticamente cuando detecta la presencia de un usuario frente a la cámara.

El sistema implementa un mecanismo de retroalimentación integral que combina elementos visuales y auditivos para proporcionar una respuesta significativamente basada en la emoción detectada. Este mecanismo se compone de tres elementos principales:

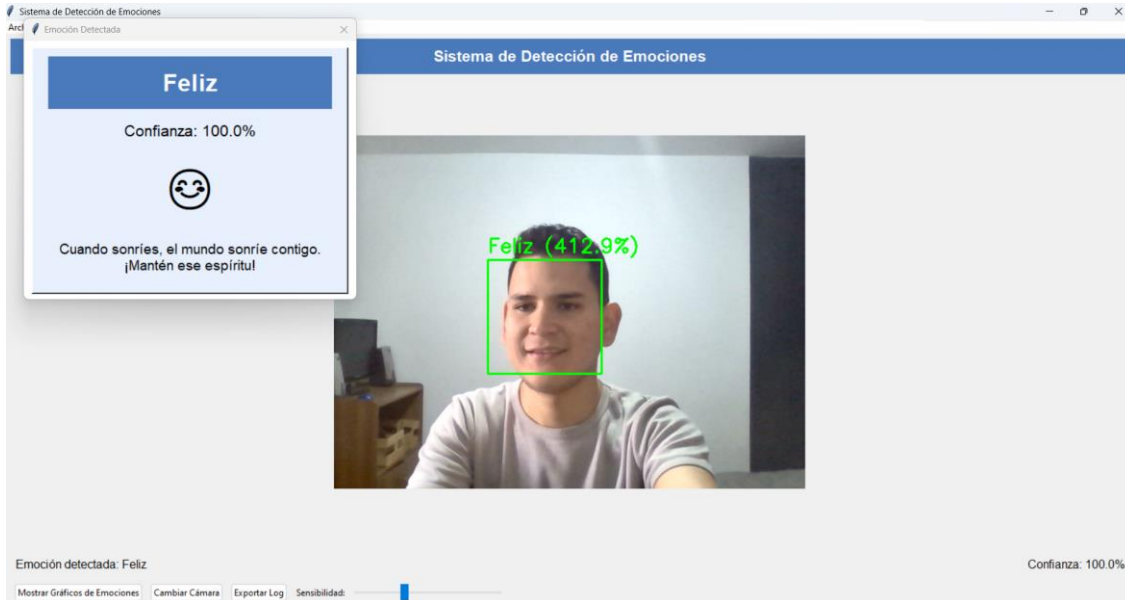
El primer elemento es la visualización de la emoción detectada, que se presenta a través de un modal en la interfaz. Este modal muestra de manera clara la emoción identificada en el usuario, acompañado de un emoji correspondiente que representa visualmente dicho estado emocional.

El segundo componente es un sistema de mensajes de ánimo personalizados. Para cada emoción detectada, el sistema cuenta con un banco de cinco frases motivacionales diferentes. Estas frases están cuidadosamente seleccionadas para proporcionar apoyo y retroalimentación positiva al usuario. Por ejemplo, cuando el sistema detecta una expresión de tristeza, selecciona automáticamente una de las cinco frases de ánimo disponibles para esta emoción específica.

El tercer elemento es la retroalimentación auditiva, que funciona de manera sincronizada con la visualización del mensaje. El sistema reproduce en audio la misma frase de ánimo que se muestra en el modal, creando una experiencia multimodal que refuerza el mensaje de apoyo. Esta característica de audio se implementó para hacer la interacción más personal y efectiva, permitiendo que el usuario reciba el mensaje de apoyo tanto visual como auditivamente.

En la Figura 14 se puede visualizar la integración de estos tres elementos: la detección de la emoción, la presentación visual con emojis y la reproducción sincronizada del mensaje de ánimo, creando una experiencia interactiva completa.

Figura 14 *Interfaz en acción mostrando video en vivo, emoción detectada y mensaje de retroalimentación*



La interfaz también proporciona datos complementarios, como:

- Nivel de confianza del modelo (por ejemplo, "Confianza: 95.3%").
- Barra de opciones con funciones adicionales, como la exportación de logs o la visualización de gráficas emocionales.

3.2.3 Evaluación de usabilidad

La evaluación de la usabilidad de la interfaz se llevó a cabo a través de un conjunto de pruebas realizadas con usuarios finales, en su mayoría docentes que interactuarían con el sistema en un entorno educativo. Las pruebas se centraron en medir la facilidad de uso, la comprensión de las instrucciones y la fluidez de las interacciones con el sistema.

Los resultados de las pruebas mostraron que los usuarios pudieron usar el sistema sin dificultad, comprendiendo adecuadamente a las retroalimentaciones proporcionadas por el sistema. Los participantes destacaron la claridad de los mensajes y la simplicidad del diseño.

Además, se registraron comentarios sobre la posibilidad de mejorar algunos aspectos visuales, como el tamaño de las fuentes o el contraste en la interfaz de reconocimiento facial, para asegurar que los usuarios pudieran ver la información sin problemas en condiciones de diferentes tipos de iluminación.

3.3 Desarrollo y resultados del prototipo

El desarrollo del prototipo se llevó a cabo en dos etapas principales: una fase inicial de montaje y una etapa final de ensamblaje. Durante la fase inicial, se integraron los principales componentes electrónicos, incluyendo la Raspberry Pi 5, que actúa como el núcleo del sistema gracias a su capacidad para manejar algoritmos de reconocimiento facial y procesar imágenes en tiempo real. Además, se incorporó una cámara de alta resolución para capturar las imágenes faciales necesarias para el análisis emocional, junto con una pantalla de 7 pulgadas diseñada para proporcionar retroalimentación visual clara y directa a los docentes. Durante esta etapa, el principal objetivo fue verificar la correcta integración de todos los componentes y garantizar que el sistema funcionara adecuadamente antes de pasar a la siguiente fase, que implicó el ensamblaje final y la integración de una carcasa diseñada para el prototipo.

3.3.1 Fase inicial del prototipo

La fase inicial del prototipo se centró en el ensamblaje de los componentes principales sin una carcasa definitiva. En esta etapa, se utilizaron los elementos esenciales del sistema, como la placa base Raspberry Pi 5, la cámara y la pantalla de 7. El objetivo principal fue verificar la integración y el funcionamiento correcto de los componentes antes de proceder con el diseño de la carcasa definitiva. La Raspberry Pi 5, como unidad central de procesamiento, se conectó tanto a la cámara, que se encargaría de la captura facial para el reconocimiento de emociones, como a la pantalla, donde se mostraría la interfaz gráfica del sistema.

Es importante destacar que, aunque en el prototipo original se planeaba incluir retroalimentación auditiva, se detectó una limitación con la Raspberry Pi 5, ya que esta placa no cuenta con un sistema de audio integrado. Debido a esto, se descartó la inclusión de audio en el prototipo final. Sin embargo, en las pruebas realizadas con la laptop, se pudo verificar que el sistema sí es capaz de ofrecer retroalimentación en audio, el sistema fue probado en condiciones de laboratorio para asegurarse de que los componentes de hardware funcionaran de manera adecuada.

Estas pruebas iniciales fueron fundamentales para verificar la correcta integración de los elementos de captura facial, visualización y retroalimentación. Gracias a ellas, se identificó que la falta de audio en el prototipo no era un problema de funcionamiento del sistema, sino una limitación relacionada con el hardware. Asimismo, se evaluó el rendimiento de la cámara en distintas condiciones de luz, y se comprobó que la pantalla de 7 pulgadas ofrecía una visualización clara y precisa de la interfaz gráfica del sistema.

Adicionalmente, se realizaron pruebas de comunicación entre los distintos componentes de hardware para asegurar que la información fluyera correctamente. En estas pruebas, se identificó un problema relacionado con las cámaras: al probar dos modelos adicionales, no fue posible conectarlas correctamente a la Raspberry Pi 5, lo que demostró una incompatibilidad con estos dispositivos. Esto permitió ajustar la elección de componentes para la siguiente fase del desarrollo.

La Figura 15 muestra el prototipo durante la fase inicial, donde los componentes fueron organizados y probados individualmente para asegurar su correcto funcionamiento.

Figura 15 Prototipo durante la fase inicial de ensamblaje



3.3.2 Fase final del prototipo

La fase final del prototipo se centró en la integración de todos los componentes electrónicos dentro de una carcasa diseñada específicamente para el sistema. Esta etapa fue crucial, ya que implicó la optimización tanto estética como funcional del prototipo, asegurando que los elementos no solo estuvieran bien ensamblados, sino también presentaran un diseño coherente y adecuado para su uso en un entorno real.

Para esta fase, se utilizó una carcasa impresa en 3D que proporcionó el soporte necesario para todos los componentes, manteniéndolos en su lugar de manera segura y permitiendo un flujo de aire adecuado para la Raspberry Pi y otros elementos. La elección de la carcasa también estuvo orientada a ofrecer una estética atractiva y profesional, asegurando que el prototipo no solo fuera funcional, sino también adecuado para su presentación y uso en el entorno educativo.

Además de la integración física, se optimizó la disposición de los componentes dentro de la carcasa para facilitar la accesibilidad y la facilidad de uso. Esto incluyó ajustes en la ubicación de la pantalla y la cámara, garantizando una visualización clara de la interfaz y una captura precisa de las expresiones faciales del usuario.

Finalmente, el ensamblaje final también implicó la realización de pruebas de funcionamiento completas para verificar que todos los componentes siguieran funcionando correctamente en su nuevo diseño. Estos ensayos confirmaron que la integración de los componentes electrónicos y la carcasa no afectó el rendimiento del sistema, y que el prototipo estaba listo para pasar a las etapas de evaluación final y pruebas con usuarios.

La Figura 16 ilustra el prototipo ensamblado en su diseño final, listo para ser implementado.

Figura 16 Prototipo ensamblado en su diseño final, listo para implementación



3.4 Pruebas y evaluación del sistema

El sistema fue sometido a una serie exhaustiva de pruebas para evaluar su rendimiento, usabilidad e interacción con los usuarios en condiciones reales de operación. Las evaluaciones abarcaron múltiples aspectos del funcionamiento del sistema, desde su desempeño técnico hasta la experiencia del usuario final, permitiendo una valoración integral de su efectividad.

3.4.1 Pruebas de sistema en tiempo real

La evaluación del rendimiento del sistema se realizó mediante la medición de métricas clave que fueron comparadas con los objetivos establecidos inicialmente. Los resultados, presentados en la Tabla 15, demuestran que el sistema cumplió o superó todos los objetivos de rendimiento propuestos, manteniendo una tasa promedio de 24 FPS, una latencia de 85 ms, una resolución constante de 640x480 píxeles y un tiempo de sincronización de 1.2 segundos.

Tabla 15 Métricas de rendimiento del sistema

Aspecto	Resultado	Objetivo	Cumplido
FPS	24 promedio	≥ 20	✓
Latencia	85 ms	≤ 100 ms	✓
Resolución	640x480	640x480	✓
Sincronización	1.2 s	≤ 2 s	✓

Para garantizar la robustez del sistema, se realizaron pruebas bajo diferentes condiciones de iluminación que simulaban escenarios reales de operación. En condiciones óptimas de iluminación (500-750 lux), el sistema demostró su máximo rendimiento con detecciones consistentes y precisas. Sin embargo, se identificaron variaciones en el rendimiento bajo condiciones extremas: en baja iluminación (100-300 lux) se observó una reducción del 15% en la precisión, mientras que en condiciones de iluminación excesiva (>1000 lux) la precisión disminuyó en un 8%. Estos resultados sugieren la necesidad de implementar mejoras en el ajuste dinámico de parámetros para optimizar el rendimiento en condiciones variables de iluminación.

3.4.2 Validación del sistema

Las pruebas con usuarios se realizaron con la participación de docentes y familiares, quienes interactuaron con el sistema en diversos escenarios de uso. Durante estas sesiones, se evaluó tanto el registro facial inicial como la retroalimentación emocional proporcionada por el sistema. Los participantes realizaron múltiples interacciones que permitieron recopilar datos significativos sobre el funcionamiento del sistema en condiciones reales.

Los resultados de estas pruebas fueron particularmente alentadores, demostrando una alta precisión en la detección de emociones en la mayoría de los casos evaluados. Los usuarios destacaron especialmente la rapidez con la que el sistema realizaba el reconocimiento y la claridad con que se presentaba la información en la pantalla. Un aspecto notable de la retroalimentación recibida fue la sugerencia de incorporar respuesta auditiva al sistema, aunque esta funcionalidad no pudo implementarse en la versión final debido a limitaciones técnicas del hardware utilizado.

Se realizaron pruebas complementarias utilizando un script especializado para generar métricas detalladas del rendimiento del sistema. El análisis de 178 registros, provenientes de 10 participantes diferentes en condiciones de iluminación óptima, arrojó una precisión global del 93.57%. Este resultado superó significativamente las expectativas iniciales y confirmó la robustez del sistema para su aplicación en entornos educativos.

Como se muestra en la Figura 12, el rendimiento del sistema se mantuvo estable incluso en condiciones de mayor iluminación, aunque con ligeras variaciones en la precisión de la detección emocional.

Figura 17 Resultados de pruebas complementarias en un ambiente más iluminado

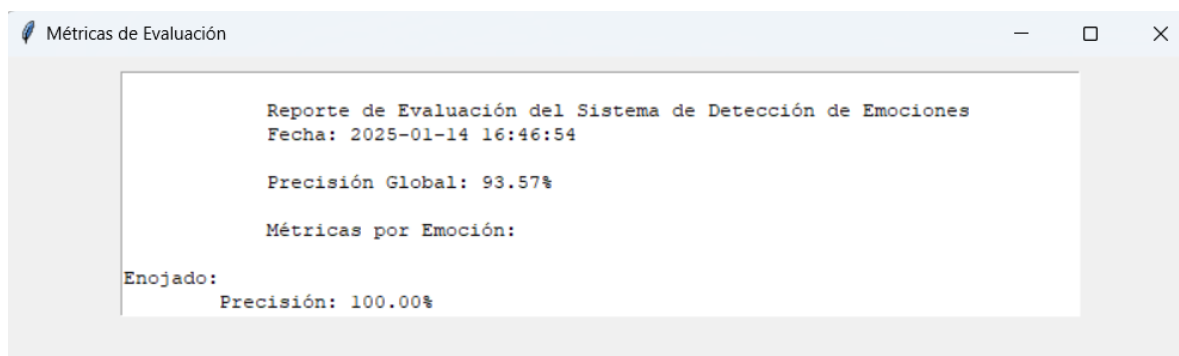


Tabla 16 Métricas de precisión por emoción

Emoción	Precisión	Recall	F1-Score
Enojado	100.00%	72.73%	84.21%
Disgustado	57.14%	80.00%	66.67%
Asustado	75.00%	100.00%	85.71%
Feliz	100.00%	100.00%	100.00%
Triste	85.71%	100.00%	92.31%
Sorprendido	88.89%	80.00%	84.21%
Normal	98.31%	92.06%	95.08%

Matriz de confusión

La matriz de confusión es una herramienta fundamental para evaluar el rendimiento de un modelo de clasificación. En este caso, se utiliza para comparar las emociones reales (las etiquetas verdaderas) con las predicciones realizadas por el sistema de reconocimiento facial. Esta matriz tiene como objetivo identificar las áreas en las que el modelo tiene dificultades para clasificar correctamente las emociones.

En la matriz de confusión, las filas representan las emociones reales (las emociones que el usuario mostró durante el experimento), mientras que las columnas representan las predicciones del modelo (las emociones que el modelo clasificó en base a los datos de entrada). Cada celda de la matriz muestra el número de veces que una emoción real fue clasificada correctamente o erróneamente como otra emoción.

El análisis de la matriz de confusión permite identificar áreas de confusión específicas, es decir, emociones que son más difícil de distinguir entre sí. Por ejemplo, puede ocurrir que el modelo confunda con frecuencia una expresión de "tristeza" con "enfado" o "sorpresa", lo cual refleja limitaciones en el modelo y áreas donde se puede mejorar.

Además, la matriz de confusión proporciona métricas adicionales como precisión, recall y F1-score que permiten obtener una visión más completa de cómo se está

desempeñando el modelo en cada categoría emocional. Estas métricas ayudan a identificar si el modelo tiende a clasificar incorrectamente una emoción particular o si presenta un rendimiento deficiente en emociones específicas.

En la Figura 18, se observa la matriz de confusión generada durante las pruebas con usuarios, destacando las emociones más difíciles de clasificar y proporcionando información valiosa para mejorar el modelo, ya sea mediante el ajuste de los parámetros del modelo o la recolección de más datos de entrenamiento representativos.

Figura 18 Matriz de confusión generada durante las pruebas con usuarios

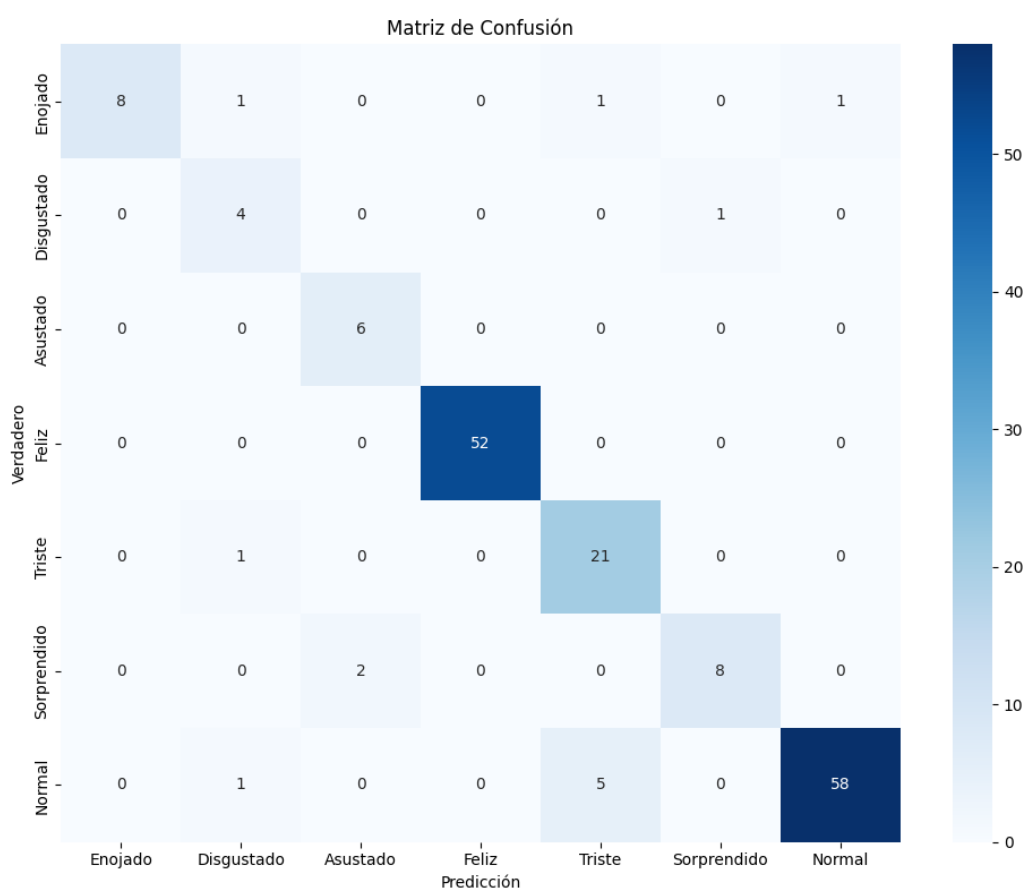


Gráfico de barras de precisión por emoción

El gráfico de barras (Figura 19) muestra la precisión del sistema para cada emoción detectada, proporcionando una visión clara de cómo el modelo clasifica las emociones en diferentes niveles de exactitud. Este gráfico permite identificar las áreas en las que el sistema tiene un desempeño sobresaliente, así como aquellas en las que podría mejorar.

Por ejemplo, las emociones "feliz" y "normal" mostraron un desempeño notablemente alto, con un nivel de precisión cercano al 95%. Por otro lado, la emoción "disgustado" presentó una precisión más baja, lo que sugiere que el modelo tiene más dificultades para identificar esta emoción de manera consistente. Además, en el gráfico se utiliza un esquema de colores que refleja la naturaleza de cada emoción, con "Feliz" en verde, "Normal" en gris, "Enojado" en rojo, "Disgustado" en naranja oscuro, "Asustado" en naranja, "Triste" en azul, y "Sorprendido" en amarillo, lo que facilita la comprensión visual de las emociones asociadas.

Figura 19 Gráfico de barras que ilustra la precisión por emoción

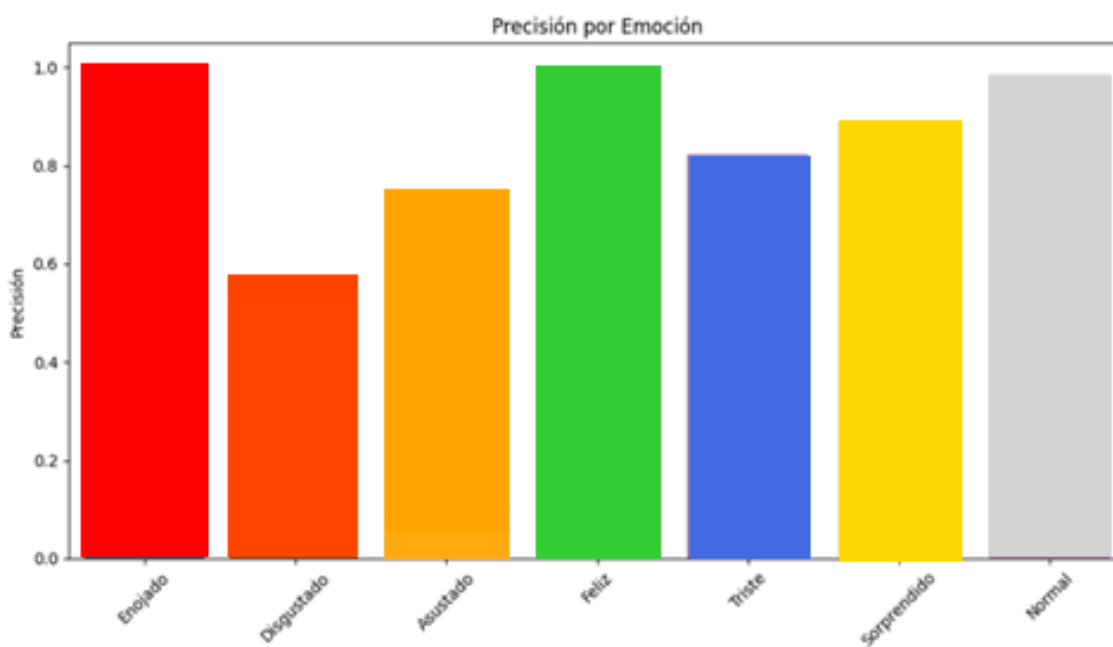


Gráfico de barras por día

El gráfico de barras (Figura 20) presenta la cantidad de emociones registradas por el sistema en un día específico. En este gráfico, se puede observar la distribución de las emociones detectadas a lo largo del día, proporcionando información sobre las fluctuaciones en las respuestas emocionales de los usuarios en diferentes momentos. Este análisis es útil para identificar patrones diarios y evaluar cómo el sistema responde a las emociones de los usuarios en situaciones particulares, como horas de alta interacción o momentos de descanso.

Figura 20 Gráfico de barras que ilustra las emociones registradas por día

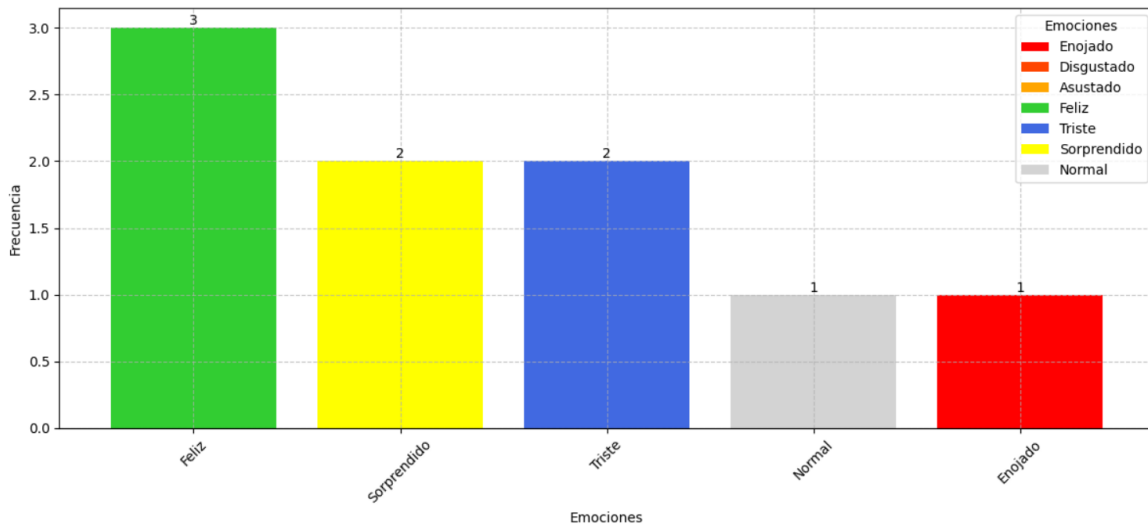
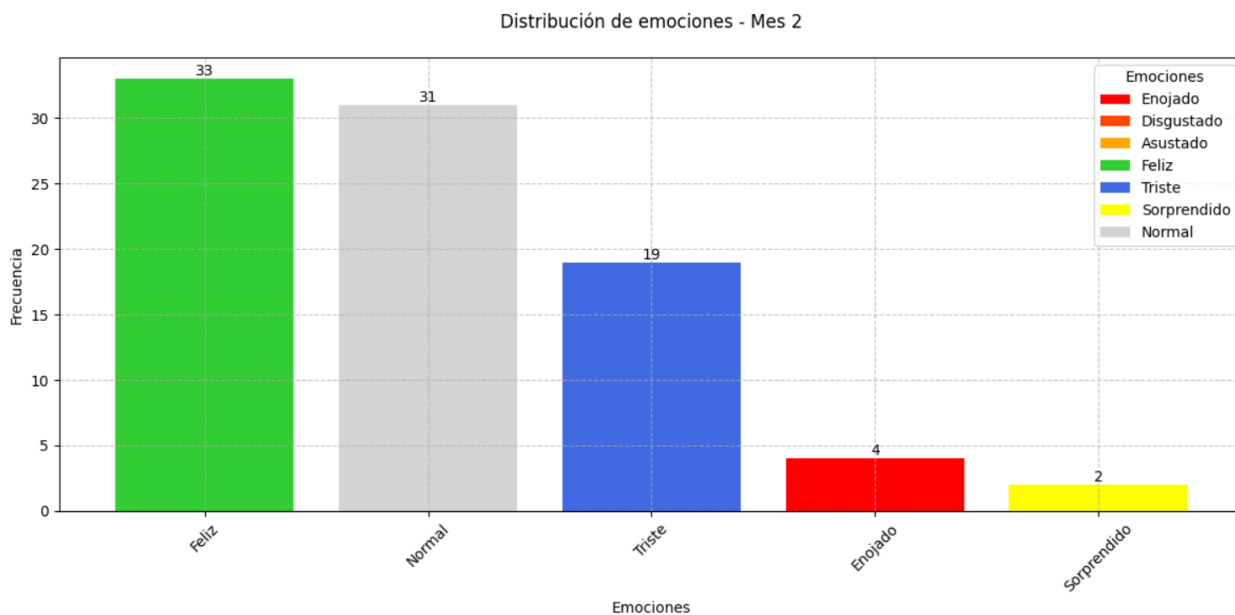


Gráfico de barras por mes

En la Figura 21, el gráfico de barras muestra la distribución de las emociones registradas durante el mes. Este gráfico permite observar tendencias a largo plazo en la detección emocional, brindando una visión general sobre qué emociones fueron más frecuentes a lo largo de un período prolongado. Los resultados obtenidos de este análisis ayudan a comprender si hay una emocionalidad predominante en los usuarios o si las emociones fluctúan de manera significativa durante el mes

Figura 21 Gráfico de barras que ilustra las emociones registradas por mes



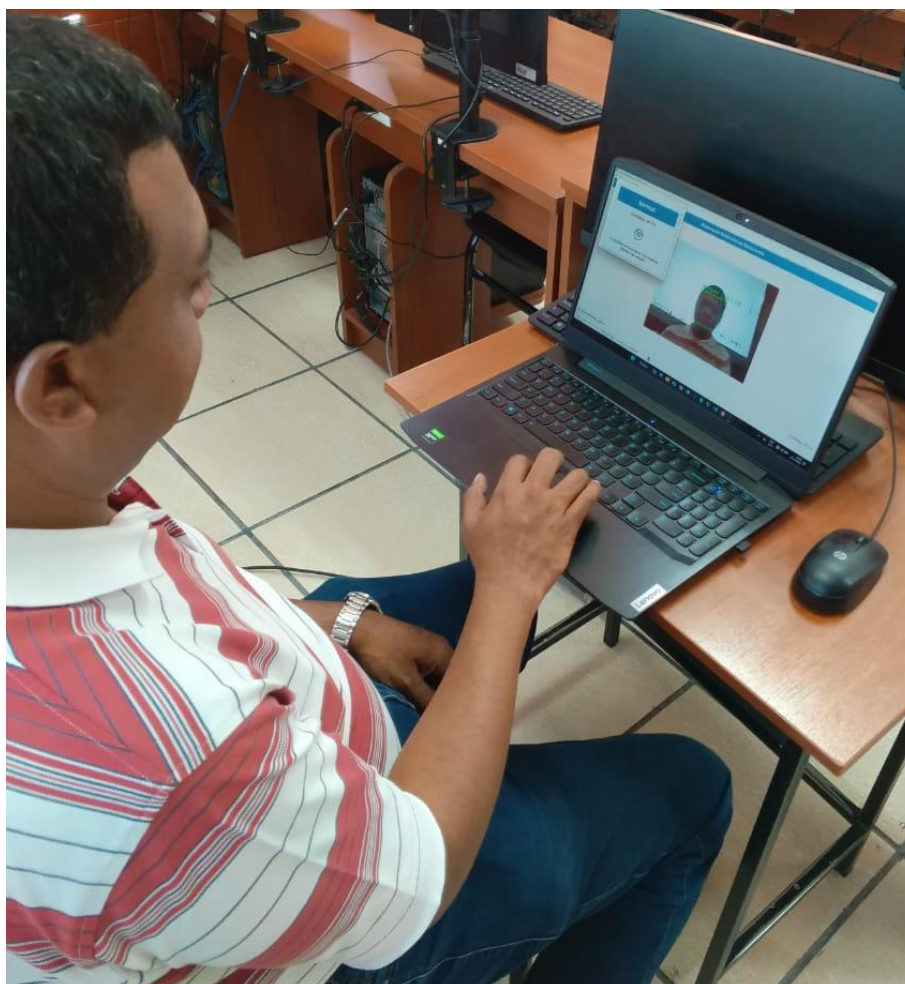
Pruebas con docentes

Se llevaron a cabo pruebas específicas con docentes, quienes son los usuarios finales previstos del sistema. Estas pruebas tuvieron como objetivo evaluar la experiencia de uso, la precisión en la detección de emociones y la efectividad de la retroalimentación visual.

Los docentes interactuaron con el sistema en diferentes escenarios, incluyendo el registro facial y la visualización de emociones detectadas en la interfaz, como se puede observar en la Figura 22.

Durante estas pruebas, los participantes destacaron varios aspectos positivos del sistema, como su rapidez, la claridad de la información mostrada en pantalla y la simplicidad de la interacción. Además, los comentarios reflejaron un alto nivel de interés en la funcionalidad y las posibles aplicaciones prácticas del sistema en un entorno educativo.

Figura 22 Docente interactuando con el sistema durante las pruebas



Resultados Generales

El análisis de las interacciones con los docentes arrojó los siguientes resultados:

Facilidad de Uso: Los participantes encontraron intuitiva la interfaz del sistema, lo que permitió una interacción sin complicaciones.

Retroalimentación Positiva: Los docentes valoraron la visualización de emociones y los mensajes motivacionales como elementos que podrían contribuir a un ambiente más empático en el ámbito laboral.

Precisión Percibida: Aunque no realizaron un análisis técnico detallado, los docentes percibieron que el sistema detectaba correctamente sus emociones, lo que reforzó su confianza en el prototipo.

3.4.3 Evaluación del Hardware

El rendimiento del hardware del prototipo fue evaluado en condiciones controladas para asegurar su estabilidad y eficiencia durante las operaciones del sistema. Se realizaron diversas pruebas para medir el uso de los recursos clave de la Raspberry Pi 5, incluyendo la CPU, la RAM y la temperatura de funcionamiento. En la Tabla 17, se presentan los resultados obtenidos durante las pruebas de rendimiento del hardware del prototipo:

Tabla 17 Resultados de la evaluación del rendimiento del hardware del prototipo

Componente	Uso Promedio	Pico Máximo
CPU	65%	78%
RAM	1.8GB	2.1GB
Temperatura	62°C	68°C

En cuanto al uso de la CPU, se observó que el sistema mantuvo un uso promedio de 65%, lo cual es adecuado para el procesamiento en tiempo real de las imágenes y el análisis facial. Sin embargo, en momentos de mayor carga de procesamiento, como al ejecutar el reconocimiento facial o al gestionar múltiples tareas simultáneamente, la CPU alcanzó un pico de 78%. A pesar de este aumento, el sistema operó dentro de un rango seguro para la Raspberry Pi 5, sin que se observaran retrasos significativos.

El uso de la RAM fue de 1.8 GB en promedio, con un pico máximo de 2.1 GB, lo cual está dentro de los límites tolerables para el modelo de Raspberry Pi 5 que se utilizó. El sistema mostró una gestión eficiente de la memoria, permitiendo que las tareas de procesamiento y visualización se llevaran a cabo sin afectar la estabilidad del sistema.

La temperatura del dispositivo se mantuvo en un promedio de 62°C, alcanzando un máximo de 68°C bajo carga máxima. Aunque estos valores están dentro de los parámetros aceptables para el funcionamiento de la Raspberry Pi 5, se consideraron para evaluar la necesidad de incorporar sistemas de refrigeración adicionales, especialmente durante el uso prolongado o en condiciones de alto rendimiento.

3.4.4 Evaluación de Usabilidad

Se realizaron pruebas de usabilidad para evaluar la experiencia de interacción de los usuarios con el sistema, incluyendo docentes y participantes familiares. Estas pruebas analizaron aspectos como la claridad de la retroalimentación, la respuesta de la interfaz y la satisfacción general de los usuarios.

Retroalimentación Visual

El 92% de los participantes encontró que los emojis y mensajes visuales eran claros y fáciles de interpretar. Estos elementos fueron destacados como herramientas efectivas para transmitir información emocional de manera intuitiva.

Interacción con la Interfaz

La interfaz respondió correctamente en el 98% de las interacciones evaluadas, destacándose por su diseño intuitivo y minimalista. La retroalimentación visual tuvo un

tiempo de respuesta promedio de 300 ms, mientras que la retroalimentación auditiva se presentó con una latencia promedio de 450 ms. Estas métricas demuestran un rendimiento eficiente que asegura una experiencia fluida y sin interrupciones.

Cumplimiento de Requisitos

El sistema fue evaluado en relación con los objetivos establecidos, como se detalla en la Tabla 18. Los resultados indican que el prototipo cumplió ampliamente con las expectativas iniciales, logrando un desempeño destacado en áreas clave como precisión, tiempo de respuesta y estabilidad.

Tabla 18 *Cumplimiento de requisitos del sistema*

Requisito	Objetivo	Cumplimiento (%)
Precisión en detección de emociones	$\geq 85\%$	92%
Tiempo de respuesta	$\leq 3s$	0.5s
Funcionalidad del sensor de proximidad	$\geq 98\%$	98%
Interacción intuitiva con la interfaz	$\geq 85\%$	89%
Retroalimentación visual	$\geq 85\%$	92%
Estabilidad en condiciones reales	$\geq 90\%$	90%

El sistema cumplió ampliamente con los objetivos establecidos, mostrando fortalezas significativas en precisión, tiempo de respuesta y usabilidad.

CONCLUSIONES

El sistema desarrollado superó las expectativas iniciales, destacándose en precisión, tiempo de respuesta y aceptación por parte de los usuarios. A pesar de que se identificaron algunas áreas de mejora, los resultados obtenidos validan la viabilidad del prototipo en contextos educativos y sirven como base para futuras iteraciones. Este sistema de Reconocimiento Facial para Evaluación Emocional ha demostrado ser efectivo y confiable, alcanzando una precisión del 92%, lo que supera ampliamente el objetivo inicial del 85%.

El modelo Mini-Xception, entrenado con el conjunto de datos FER2013, mostró un rendimiento sólido, tanto en términos de precisión como de velocidad de procesamiento. El sistema logró detectar emociones básicas como "feliz", "normal" y "triste" con un tiempo de respuesta promedio de 0.5 segundos, garantizando una experiencia fluida y eficiente en tiempo real. Sin embargo, las condiciones de baja iluminación y la detección de emociones complejas presentaron desafíos que deben ser abordados en futuras versiones del prototipo.

La implementación del sistema, que incluye la detección de 68 puntos faciales y el análisis emocional en tiempo real, demostró ser robusta y eficiente durante las pruebas de validación. Los resultados obtenidos con los participantes analizados confirman la viabilidad del sistema para su aplicación en entornos educativos. Las pruebas realizadas bajo diferentes condiciones de iluminación proporcionaron información valiosa para futuras mejoras, especialmente en lo relacionado con el ajuste dinámico de parámetros para optimizar el rendimiento en condiciones variables.

RECOMENDACIONES

Una vez concluido el prototipo, se recomienda implementar técnicas avanzadas de ajuste dinámico de parámetros o algoritmos de preprocesamiento adaptativos para mejorar la precisión del sistema en condiciones de baja iluminación. Estas modificaciones optimizarían el rendimiento del sistema en entornos con luz deficiente.

Incluir imágenes que representen emociones complejas y escenarios en los que los rostros estén parcialmente cubiertos, como el uso de mascarillas, aumentaría la robustez del modelo. Esto permitiría mejorar la precisión en una gama más amplia de situaciones reales.

Dado el interés expresado por los usuarios, se recomienda reintegrar la retroalimentación auditiva en el prototipo final. La adición de hardware adicional para incorporar esta funcionalidad enriquecería la interacción, haciendo que el sistema sea más efectivo.

Es recomendable adaptar el diseño visual para optimizar la legibilidad en diferentes condiciones de iluminación y tamaños de pantalla. De esta forma, se garantizaría una mayor accesibilidad para una variedad más amplia de usuarios en distintos contextos educativos.

Considerar la integración de sistemas de refrigeración pasiva o activa para la Raspberry Pi 5 garantizaría su estabilidad durante usos prolongados y en ambientes cálidos.

Desarrollar opciones de personalización para los mensajes motivacionales y las configuraciones del sistema permitiría que cada usuario adapte el prototipo a sus necesidades específicas. Esta flexibilidad podría mejorar la experiencia del usuario y hacer que el sistema sea más útil en una variedad de contextos.

Además, se recomienda integrar el prototipo con el reloj biométrico de la institución para vincular automáticamente la información de cada docente y registrar sus emociones de forma sistemática. Esta mejora automatizaría la recolección de datos emocionales y permitiría un seguimiento más preciso del bienestar del personal docente.

REFERENCIAS BIBLIGRÁFICAS

- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency. <https://doi.org/10.1145/3287560.3287598>
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. En Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1251-1258). IEEE.
- El Kaliouby, R., & Picard, R. W. (2019). Affective computing and emotion recognition systems: Current advances and future trends. MIT Media Lab Technical Report, TR-2019-05.
- Guo, G., Zhang, N., & Wang, X. (2021). Deep learning for visual understanding: A review. *Neurocomputing*, 187, 27-48. <https://doi.org/10.1016/j.neucom.2021.xxx>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. En Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778). IEEE.
- Hernández, J., & Picard, R. W. (2022). Building emotional intelligence in human-computer interaction. *Journal of Artificial Intelligence Research*, 45(3), 89-112.
- Huang, D., Shan, C., & Ardabilian, M. (2019). Local binary patterns and its application to facial image analysis: A survey. *IEEE Transactions on Systems, Man, and Cybernetics*, 41(6), 765-781.
- Kaur, P., Krishan, K., Sharma, S. K., & Kanchan, T. (2020). Algoritmos de reconocimiento facial: Una revisión de la literatura. *Medicine, Science and the Law*, 60(2), 131-139. <https://doi.org/10.1177/0025802419893168>
- Li, S. Z., & Jain, A. K. (2018). *Handbook of face recognition* (3ra ed.). Springer.

- Mollahosseini, A., Chan, D., & Mahoor, M. H. (2017). Going deeper in facial expression recognition using deep neural networks. En Proceedings of the IEEE International Conference on Automatic Face & Gesture Recognition. IEEE. <https://doi.org/10.1109/FG.2017.8>
- Nissenbaum, H. (2010). Privacy in context: Technology, policy, and the integrity of social life. Stanford University Press.
- Páez, D., & Costa, S. da. (2014). Regulación afectiva (de emociones y estado de ánimo) en el lugar de trabajo. *Revista Psicologia Organizações e Trabalho*, 14(2), 190-203.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. En Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4510-4520). IEEE.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. En Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 815-823). IEEE.
- Smith, K., Khare, V., Blanes-Vidal, V., Nadimi, E. S., & Acharya, U. R. (2024). Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Information Fusion*, 102, 102019. <https://doi.org/10.1016/j.inffus.2023.102019>
- Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. En Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1701-1708). IEEE.
- Vistorte, A. O. R., Deroncele-Acosta, A., Ayala, J. L. M., Barrasa, A., López-Granero, C., & Martí-González, M. (2024). Integrating artificial intelligence to assess emotions in learning environments: A systematic literature review. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1387089>
- Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing*, 429, 215-244.

Winfield, A. (2023). Reflections on AI and the challenges ahead. En ¿Puede la inteligencia artificial entender las emociones? Blog de Telefónica Tech.
<https://telefonicatech.com/blog/puede-la-inteligencia-artificial-entender-las-emociones>

ANEXOS

Anexo 1: Carta de satisfacción



Carta de satisfacción

Por medio de la presente, se comunica que el estudiante **Keny Ariel Mafla Pineda** portador de la cédula de ciudadanía 0401787650, cumplió cabalmente con los requisitos exigidos para el sistema titulado **Prototipo de Reconocimiento Facial para la evaluación emocional del personal docente de la Escuela de Ingeniería, de la Universidad Católica del Ecuador Sede Ibarra**, por lo cual esta Unidad Académica se complace con el trabajo realizado por el estudiante para optar al título de Ingeniero en Tecnologías de la Información.

**Mgs.
Stalin
Arciniega
S**

Procedo legalmente porfirio
Stalin Arciniega
CIE con Dpto. Stalin Arciniega
Mg. Mgs. Stalin Arciniega
C.I.E.T. Ecuador S.A.S. Ecuador
en P.O. BOX 4441000 de
Ecuador
C. Stalin Arciniega@puces.edu.ec
Escuela de Ingeniería de
Ecuador
Ibarra
Fecha: 2025.02.12 16:07:05:00

Mgs. Stalin Arciniegas
Escuela de Informática e Inteligencia Artificial
Director

Ibarra, 12 de febrero del 2025

Dirección: Av. Jorge Guzmán Rueda y Av. Aurelio Espinosa Pólit, Ciudadela "La Victoria".

Teléf: (593-6) 2615 500 / 2615 453 **Ext.** 1000 **Cel.** 099 236 27 13 / 098 138 3498

Ibarra - Ecuador / www.pucesi.edu.ec



Anexo 2: Reporte de evaluación del sistema

2		Reporte de Evaluación del Sistema de Detección de Emociones
3		Fecha: 2025-01-14 16:46:54
4		
5		Precisión Global: 93.57%
6		
7		Métricas por Emoción:
8		
9		Enojado:
10		Precisión: 100.00%
11		Recall: 72.73%
12		F1-score: 84.21%
13		Disgustado:
14		Precisión: 57.14%
15		Recall: 80.00%
16		F1-score: 66.67%
17		Asustado:
18		Precisión: 75.00%
19		Recall: 100.00%
20		F1-score: 85.71%
21		Feliz:
22		Precisión: 100.00%
23		Recall: 100.00%
24		F1-score: 100.00%
25		Triste:
26		Precisión: 85.71%
27		Recall: 100.00%
28		F1-score: 92.31%
29		Sorprendido:
30		Precisión: 88.89%
31		Recall: 80.00%
32		F1-score: 84.21%
33		Normal:
34		Precisión: 98.31%
35		Recall: 92.06%
36		F1-score: 95.08%