



**UNIVERSIDAD TÉCNICA DE AMBATO**

**FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E  
INDUSTRIAL**

**CARRERA DE TELECOMUNICACIONES**

**Tema:**

---

**SISTEMA DE DETECCIÓN DE EMOCIONES MEDIANTE EL ANÁLISIS  
DE INDICADORES FACIALES EMPLEANDO INTELIGENCIA  
ARTIFICIAL**

---

Trabajo de titulación modalidad Proyecto de Investigación, presentado previo a la  
obtención del título de Ingeniero en Telecomunicaciones

**ÁREA:** Comunicaciones

**LÍNEA DE INVESTIGACIÓN:** Tecnologías de la información y Sistemas de  
control

**AUTOR:** Daniel Rodrigo Fiallos Valladares

**TUTOR:** Ing. Edgar Patricio Córdova Córdova, Mg.

**Ambato - Ecuador**

**febrero – 2024**

## **APROBACIÓN DEL TUTOR**

En calidad de tutor del trabajo de titulación con el tema: SISTEMA DE DETECCIÓN DE EMOCIONES MEDIANTE EL ANÁLISIS DE INDICADORES FACIALES EMPLEANDO INTELIGENCIA ARTIFICIAL, desarrollado bajo la modalidad Proyecto de Investigación por el señor Daniel Rodrigo Fiallos Valladares, estudiante de la Carrera de Telecomunicaciones, de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial, de la Universidad Técnica de Ambato, me permito indicar que el estudiante ha sido tutorado durante todo el desarrollo del trabajo hasta su conclusión, de acuerdo a lo dispuesto en el Artículo 17 del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato y el numeral 6.3 del instructivo del reglamento referido.

Ambato, febrero 2024.

-----  
Ing. Edgar Patricio Córdova Córdova, Mg.

TUTOR

## AUTORÍA

El presente trabajo de titulación con el tema: SISTEMA DE DETECCIÓN DE EMOCIONES MEDIANTE EL ANÁLISIS DE INDICADORES FACIALES EMPLEANDO INTELIGENCIA ARTIFICIAL es absolutamente original, auténtico y personal y ha observado los preceptos establecidos en la Disposición General Quinta del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato. En tal virtud, el contenido, efectos legales y académicos que se desprenden del mismo son de exclusiva responsabilidad del autor.

Ambato, febrero 2024.



---

Daniel Rodrigo Fiallos Valladares

C.C. 1804420816

AUTOR

## DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato para que reproduzca total o parcialmente este trabajo de titulación dentro de las regulaciones legales e institucionales correspondientes. Además, cedo todos mis derechos de autor a favor de la institución con el propósito de su difusión pública, por lo tanto, autorizo su publicación en el repositorio virtual institucional como un documento disponible para la lectura y uso con fines académicos e investigativos de acuerdo con la Disposición General Cuarta del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato.

Ambato, febrero 2024.



---

Daniel Rodrigo Fiallos Valladares

C.C. 1804420816

AUTOR

## **APROBACIÓN DEL TRIBUNAL DE GRADO**

En calidad de par calificador del informe final del trabajo de titulación presentado por el señor Daniel Rodrigo Fiallos Valladares, estudiante de la Carrera de Telecomunicaciones, de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial, bajo la Modalidad Proyecto de Investigación, titulado SISTEMA DE DETECCIÓN DE EMOCIONES MEDIANTE EL ANÁLISIS DE INDICADORES FACIALES EMPLEANDO INTELIGENCIA ARTIFICIAL, nos permitimos informar que el trabajo ha sido revisado y calificado de acuerdo al Artículo 19 del Reglamento para la Titulación de Grado en la Universidad Técnica de Ambato y el numeral 6.4 del instructivo del reglamento referido. Para cuya constancia suscribimos, conjuntamente con la señora Presidente del Tribunal.

Ambato, febrero 2024.

-----  
Ing. Elsa Pilar Urrutia Urrutia, Mg.  
PRESIDENTE DEL TRIBUNAL

-----  
Ing. Jesús Guamán Molina, Mg.  
PROFESOR CALIFICADOR

-----  
Ing. Fabián Salazar Escobar, PhD.  
PROFESOR CALIFICADOR

## DEDICATORIA

*Este proyecto de investigación va dedicado a toda mi familia que son el pilar fundamental de mi vida.*

*A mi madre Edith mi eterna fuente de inspiración. Gracias por tu amor incondicional, paciencia y por ser mi mayor motivación. Este logro es tan tuyo como mío.*

*A mi padre Luis y a mi hermano Fernando, quienes han demostrado una fe inquebrantable en mí y me han brindado todo su apoyo y cariño incondicional en este camino.*

***Daniel Fiallos***

## AGRADECIMIENTO

*Expreso mi sincero agradecimiento a Dios, mi guía constante y protector, por otorgarme la fuerza y sabiduría necesarias para cumplir esta meta.*

*Agradezco profundamente a mi familia por el apoyo y aliento que me han brindado, pueden estar seguros de que les atribuiré todo ese respaldo en mí.*

*A Juan x3, Will, Isa, Erick, Danna y Palan, gracias por sus ánimos, apoyo y momentos de distracción, sin ustedes, este camino no habría sido tan magnifico, este logro también les pertenece. ¡A seguir celebrando éxitos y alegrías juntos!*

*A Shirley, por el inmenso apoyo y cariño inquebrantable, tu presencia, paciencia y aliento han sido fundamentales para mantenerme enfocado y motivado durante los momentos más desafiantes de mi vida.*

*A mi tutor de tesis Ing. Patricio Cordova, por el tiempo dedicado a apoyarme y guiarme para alcanzar este logro, su orientación ha sido fundamental en este proceso.*

**Daniel Fiallos**

## ÍNDICE GENERAL DE CONTENIDOS

<b>PORTADA</b> .....	<b>i</b>
<b>APROBACIÓN DEL TUTOR</b> .....	<b>ii</b>
<b>AUTORÍA</b> .....	<b>iii</b>
<b>DERECHOS DE AUTOR</b> .....	<b>iv</b>
<b>APROBACIÓN DEL TRIBUNAL DE GRADO</b> .....	<b>v</b>
<b>DEDICATORIA</b> .....	<b>vi</b>
<b>AGRADECIMIENTO</b> .....	<b>vii</b>
<b>ÍNDICE GENERAL DE CONTENIDOS</b> .....	<b>viii</b>
<b>ÍNDICE DE TABLAS</b> .....	<b>xi</b>
<b>ÍNDICE DE FIGURAS</b> .....	<b>xiii</b>
<b>ÍNDICE DE ANEXOS</b> .....	<b>xviii</b>
<b>RESUMEN EJECUTIVO</b> .....	<b>xix</b>
<b>ABSTRACT</b> .....	<b>xx</b>
<b>CAPÍTULO I. MARCO TEÓRICO</b> .....	<b>1</b>
1.1 Tema de investigación .....	1
1.1.1 Planteamiento del problema .....	1
1.2 Antecedentes investigativos .....	3
1.3 Fundamentación teórica .....	6



1.3.1 Emociones .....	6
1.3.2 Características del rostro .....	8
1.3.3 Indicadores faciales para la detección de emociones .....	9
1.3.4 Procesamiento digital de Señales .....	10
1.3.5 Inteligencia Artificial .....	14
1.3.6 Machine Learning .....	19
1.3.7 Redes Neuronales Artificiales.....	21
1.3.8 Redes Neuronales Convolucionales.....	24
1.3.9 Modelos de inteligencia artificial.....	27
1.4 Objetivos .....	32
1.4.1 Objetivo general.....	32
1.4.2 Objetivos específicos .....	32
<b>CAPÍTULO II. METODOLOGÍA .....</b>	<b>33</b>
2.1 Materiales.....	33
2.2 Métodos.....	33
2.2.1 Modalidad de la investigación .....	33
2.2.2 Población y muestra .....	34
2.2.3 Recolección de información.....	34
2.2.4 Procesamiento y análisis de datos .....	34
<b>CAPÍTULO III. RESULTADOS Y DISCUSIÓN.....</b>	<b>36</b>
3.1 Análisis y discusión de los resultados .....	36

3.2 Desarrollo de la propuesta.....	36
3.2.1 Emociones definidas .....	37
3.2.2 Tecnologías utilizadas.....	42
3.2.3 Arquitectura del sistema.....	57
3.2.4 Diseño del sistema.....	60
3.2.5 Pruebas de funcionamiento .....	95
3.2.6 Resultados .....	101
3.2.7 Validación de resultados .....	108
3.2.8 Presupuesto .....	117
<b>CAPÍTULO IV. CONCLUSIONES Y RECOMENDACIONES .....</b>	<b>119</b>
a. Conclusiones .....	119
b. Recomendaciones.....	120
<b>REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>122</b>
<b>ANEXOS .....</b>	<b>130</b>

## ÍNDICE DE TABLAS

Tabla 1. Emociones principales. ....	7
Tabla 2. Emociones secundarias. ....	8
Tabla 3. Atributos principales de la detección del rostro.....	8
Tabla 4. Movimientos principales de los músculos en las expresiones faciales.....	10
Tabla 5. Características del procesamiento digital de imágenes.....	12
Tabla 6. Aplicaciones del procesamiento digital de imágenes. ....	13
Tabla 7. Procedimientos del procesamiento de imágenes con IA.....	14
Tabla 8. Componentes faciales importantes.....	18
Tabla 9. Patrones faciales de las emociones. ....	19
Tabla 10. Tipos de aprendizaje de automático.....	20
Tabla 11. Algoritmos de aprendizaje automático. ....	21
Tabla 12. Ventajas de una Red Neuronal Artificial.....	22
Tabla 13. Procesos de construcción de modelos.....	27
Tabla 14. Categorías de la matriz de confusión. ....	30
Tabla 15. Métricas de rendimiento del modelo.....	31
Tabla 16. Recolección de información.....	34
Tabla 17. Emociones propuestas.....	37
Tabla 18. Problemas en la detección de elementos faciales.....	38
Tabla 19. Indicadores faciales de las emociones. ....	39
Tabla 20. Características de las emociones definidas.....	40

Tabla 21. Comparación de placas SBC.....	44
Tabla 22. Comparación de las cámaras.....	47
Tabla 23. Software de programación. ....	48
Tabla 24. Comparación de software de programación.....	49
Tabla 25. Especificaciones de los algoritmos para la red neuronal. ....	52
Tabla 26. Comparación de Web Hosting. ....	54
Tabla 27. Comparación de herramientas de entrenamiento.....	56
Tabla 28. Bibliotecas requeridas del sistema .....	61
Tabla 29. Matriz de confusión. ....	102
Tabla 30. Matriz de confusión. ....	103
Tabla 31. Matriz de confusión. ....	103
Tabla 32. Matriz de confusión. ....	104
Tabla 33. Matriz de confusión. ....	105
Tabla 34. Matriz de confusión. ....	106
Tabla 35. Matriz de confusión del disgusto. ....	107
Tabla 36. Resultados generales de la pruebas del sistema. ....	116
Tabla 37. Presupuesto de la implementación del proyecto. ....	118

## ÍNDICE DE FIGURAS

Figura 1. Las emociones en las personas. ....	6
Figura 2. Rueda de Emociones de Ekman. ....	7
Figura 3. Indicadores faciales. ....	9
Figura 4. Sistema Digital de Adquisición de Datos. ....	11
Figura 5. Procesamiento digital de imágenes. ....	11
Figura 6. Diagrama básico de la visión artificial. ....	15
Figura 7. Diagrama básico del Reconocimiento facial. ....	16
Figura 8. Proceso de reconocimiento facial. ....	17
Figura 9. Puntos de referencia del rostro. ....	18
Figura 10. Machine Learning. ....	19
Figura 11. Aprendizaje supervisado vs no supervisado. ....	21
Figura 12. Arquitectura de una Red Neuronal. ....	23
Figura 13. Características de las capas de una red neuronal artificial. ....	23
Figura 14. Diagrama de una Red Neuronal Convolutacional. ....	24
Figura 15. Arquitectura de TensorFlow. ....	25
Figura 16. Características de Keras en Deep Learning. ....	26
Figura 17. Atributos preferibles de nuestro modelo. ....	29
Figura 18. Matriz de confusión. ....	30
Figura 19. Raspberry Pi 4. ....	42

Figura 20. Jetson Nano.....	43
Figura 21. Coral Dev Board.....	44
Figura 22. Cámara web Gazal Sx1023.....	45
Figura 23. Cámara HBVCAM-HPLCC .....	46
Figura 24. Cámara web Logitech C270 HD 30FPS.....	47
Figura 25. Arquitectura del sistema. ....	58
Figura 26. Diagrama de bloques del sistema .....	60
Figura 27. Diagrama de flujo de la etapa de adquisición.....	62
Figura 28. Diagrama de flujo del procesamiento de imágenes.....	64
Figura 29. Resultados de la instalación de la biblioteca TensorFlow.....	66
Figura 30. Arquitectura de la base de datos.....	67
Figura 31. Proceso de la descompresión del data set.....	69
Figura 32. Resultados de la extracción del data set.....	69
Figura 33. Resultados del aumento de datos.....	70
Figura 34. Imágenes de entrenamiento y validación.....	71
Figura 35.Arquitectura de ResNet.....	72
Figura 36. Capas del algoritmo .....	73
Figura 37. Compilación del modelo de la red neuronal.....	75
Figura 38. Datos almacenados de entrada.....	76
Figura 39. Guardado de datos.....	77
Figura 40. Distribución normal.....	77

Figura 41. Matriz de correlación.....	78
Figura 42. Valores de los pesos.....	79
Figura 43. Representación gráfica de los pesos. ....	79
Figura 44. Resultados del entrenamiento con 23 épocas. ....	80
Figura 45. Resultados del entrenamiento con 50 épocas. ....	80
Figura 46. Entorno de TensorBoard.....	81
Figura 47. Precisión del modelo con 23 épocas. ....	82
Figura 48. Precisión del modelo con 50 épocas. ....	82
Figura 49. Pérdidas en el entrenamiento con 23 épocas. ....	83
Figura 50. Pérdidas en el entrenamiento con 50 épocas. ....	83
Figura 51. Matriz de confusión del modelo con 23 épocas.....	83
Figura 52. Matriz de confusión del modelo con 50 épocas.....	84
Figura 53. Resultados del modelo con 23 épocas ....	84
Figura 54. Resultados del modelo con 50 épocas ....	85
Figura 55. Ingreso al panel de 000web host.....	85
Figura 56. Administración de la base de datos. ....	86
Figura 57. Entorno de base de datos. ....	86
Figura 58. Representación de la base de datos.....	87
Figura 59. Apartados de la base de datos.....	87
Figura 60. Base de datos de los usuarios.....	88
Figura 61. Posicionamiento del usuario con el sistema. ....	89

Figura 62. Inicio de sesión en la interfaz de usuario.....	89
Figura 63. Lista de test creados para determinar las emociones.....	90
Figura 64. Interfaz de activación de cámara web.....	91
Figura 65. Activación del sistema.....	91
Figura 66. Ingreso a la interfaz administrativa.....	92
Figura 67. Visualización de la opción exámenes en la interfaz administrativa. ....	92
Figura 68. Visualización de la opción sesiones en la interfaz administrativa.....	93
Figura 69. Visualización de los patrones de emociones del usuario.....	93
Figura 70. Lista de usuarios activos.....	94
Figura 71. Historial etiquetado de las emociones. ....	94
Figura 72. Formulario para crear usuarios.....	95
Figura 73. Activación del usuario.....	96
Figura 74. Emoción de felicidad detectada.....	96
Figura 75. Emoción de enojo detectada.....	97
Figura 76. Emoción neutra detectada.....	97
Figura 77. Emoción de sorpresa detectada.....	97
Figura 78. Patrón de emociones del usuario. ....	98
Figura 79. Activación del usuario.....	98
Figura 80. Emoción neutra detectada.....	99
Figura 81. Emoción de felicidad detectada.....	99
Figura 82. Emoción de tristeza detectada.....	99



Figura 83. Patrón de emociones del usuario. ....	100
Figura 84. Activación del usuario. ....	100
Figura 85. Emoción neutro detectada.....	101
Figura 86. Patrón de emociones del usuario. ....	101
Figura 87. Categorización variable 1. ....	130
Figura 88. Categorización variable 2 .....	130
Figura 89. Constelación de ideas variable 1.....	131
Figura 90. Constelación de ideas variable 2.....	131
Figura 91. Data sheet de la Jetson Nano. ....	132
Figura 92. Especificaciones Técnicas de la Jetson Nano. ....	133
Figura 93. Jetson Nano Adquirida .....	133
Figura 94. Implementación del sistema. ....	134
Figura 95. Sistema en un entorno real.....	134

## ÍNDICE DE ANEXOS

Anexo A. Categorización de variables.....	130
Anexo B. NVIDIA JETSON NANO MODULE .....	132
Anexo C. Implementación del sistema.....	134
Anexo D. Implementación del sistema .....	135

## RESUMEN EJECUTIVO

Las emociones son esenciales en varios ámbitos de la vida, sin embargo, su detección y comprensión pueden llegar a ser complicadas, esto lleva a malentendidos y obstaculiza la comunicación, impactando de manera negativa en las relaciones sociales de las personas. En este contexto, se ha llevado a cabo la investigación con el objetivo de implementar un sistema de detección de emociones mediante el análisis de indicadores faciales empleando inteligencia artificial y que permita visualizar las emociones que puede tener una persona durante un tiempo determinado. El sistema se encuentra dividido en tres etapas, inicia con la adquisición y procesamiento de datos mediante la activación y el uso de una cámara web, respaldada por la biblioteca OpenCV para las técnicas de procesamiento de imágenes. La fase de entrenamiento implica el desarrollo de un modelo de aprendizaje profundo empleando Redes Neuronales Convolucionales a partir del reconocimiento facial utilizando FaceNet, perfeccionado su diseño a través del ajuste de datos, la arquitectura de la red neuronal se ha enfocado en la extracción y aprendizaje de características relevantes. Finalmente, en la etapa de almacenamiento y visualización, los datos son procesados por la Jetson Nano y se envían a un entorno web hosting que recibe los resultados y son transmitidos a la interfaz administrativa para la gestión y visualización del reporte de emociones del usuario. Los resultados de las pruebas indicaron que el sistema capturó frames cada 4 segundos, y cuenta con una precisión de clasificación del 92%, considerando que el modelo presenta una habilidad destacada para clasificar las emociones en tiempo real.

**Palabras clave:** Indicadores faciales, redes neuronales, inteligencia artificial.

## ABSTRACT

Emotions are essential in several areas of life, however, their detection and understanding can become complicated, this will lead to misunderstandings and hinder communication, negatively impacting people's social relationships. In this context, the research was carried out with the aim of implementing an emotion detection system by analyzing facial indicators using artificial intelligence and visualizing the emotions that a person may have for a period of time. The system is divided into three stages, starting with the acquisition and processing of data through the activation and use of a webcam, supported by the OpenCV library for image processing techniques. The training phase involves the development of a deep learning model using Convolutional Neural Networks from facial recognition using FaceNet, perfected its design through data fitting, the architecture of the neural network focused on the extraction and learning of relevant features. Finally, the storage and visualization stage, the data is processed by the Jetson Nano and sent to a web hosting environment that receives the results and transmits them to the administrative interface for the management and visualization of the user's emotion report. The test results indicated that the system captured frames every 4 seconds, and boasts a classification accuracy of 92%, considering that the model has an outstanding ability to classify emotions in real time.

**Keywords:** Facial indicators, neural networks, artificial intelligence.

# CAPÍTULO I. MARCO TEÓRICO

## 1.1 Tema de investigación

### SISTEMA DE DETECCIÓN DE EMOCIONES MEDIANTE EL ANÁLISIS DE INDICADORES FACIALES EMPLEANDO INTELIGENCIA ARTIFICIAL

#### 1.1.1 Planteamiento del problema

Las emociones son importantes en diversos aspectos de la vida, como en las interacciones personales, el ámbito laboral y el bienestar mental, sin embargo, su detección y comprensión precisas suelen ser complicadas. Las diferencias en cómo las personas pueden llegar a experimentar y expresar emociones llevan a malentendidos y obstaculización de la comunicación. La falta de una herramienta confiable para identificar y comprender emociones en tiempo real limita la respuesta efectiva a las necesidades emocionales en diversos contextos, que van desde el servicio al cliente y salud mental hasta la toma de decisiones en la industria [1].

A nivel mundial, la detección y comprensión de las emociones humanas se han convertido en un campo de investigación y desarrollo de gran relevancia. La mayoría de los estudios delimitan los conceptos y las características principales de cada emoción además de sintetizar los hallazgos científicos que avalan la influencia de las emociones negativas clásicas en el proceso salud-enfermedad. El avance de la tecnología ha hecho que se creen aplicaciones que aborden las emociones humanas y ayuden en sectores como la salud mental, la interacción humano-computadora y la atención al cliente impulsando a varias personas y empresas el interés en esta tecnología [2].

En América Latina, la detección y comprensión de las emociones humanas adquiere una relevancia significativa en diversos ámbitos de la sociedad. Según datos recientes, la región enfrenta desafíos particulares relacionados con la salud mental, donde se estima que aproximadamente el 20% de la población experimenta trastornos emocionales en algún momento de sus vidas. Además, en el ámbito educativo, se ha observado que el 40% de los estudiantes de la región informan niveles significativos

de estrés y ansiedad relacionados con las demandas académicas. En el sector de la atención al cliente, un estudio revela que más del 60% de los consumidores latinoamericanos espera una experiencia personalizada y emocionalmente satisfactoria al interactuar con empresas y servicios. Las diferencias culturales y la falta de recursos en algunas áreas pueden influir en la eficacia y la aplicabilidad de estos sistemas. Además, las preocupaciones éticas relacionadas con la privacidad y el acceso a la tecnología deben abordarse para garantizar su implementación responsable en la región [3].

En el contexto de Ecuador, la detección de emociones mediante análisis de indicadores faciales representa un campo de investigación y aplicación emergente. A medida que el país avanza hacia una mayor digitalización en sectores como la educación y la atención médica, se exploran oportunidades para utilizar esta tecnología en beneficio de la sociedad. Sin embargo, se presentan desafíos específicos, como la necesidad de adaptar los sistemas a la diversidad cultural de la población ecuatoriana y garantizar que el acceso a la tecnología sea equitativo en un contexto donde las disparidades económicas y de acceso a la educación son una realidad. El abordaje de estas cuestiones es esencial para asegurar que la detección de emociones basada en indicadores faciales contribuya de manera positiva al bienestar y al progreso del país [4].

En diversas esferas de la sociedad, la detección precisa de las emociones es un desafío crítico que afecta la toma de decisiones y las interacciones humanas. En el contexto de la salud mental, la dificultad para identificar emociones conduce a tratamientos inadecuados, lo que puede prolongar el sufrimiento de los pacientes. De manera similar, en el ámbito comercial, la incapacidad para comprender las emociones de los clientes resulta en insatisfacción y decisiones de compra que no cumplen las expectativas. Finalmente, las interacciones personales se ven afectadas por una comunicación deficiente debido a la falta de detección emocional, lo que puede generar tensiones y conflictos [1].

## 1.2 Antecedentes investigativos

La investigación realizada es resultado de la recolección de varias tesis y artículos de investigación referentes y previos al tema “sistema de detección de emociones mediante el análisis de indicadores faciales empleando inteligencia artificial”, enfocándose principalmente en fuentes tecnológicas que pueden proporcionar orientación en la investigación, se incluyen referencias procedentes de repositorios universitarios y bibliotecas virtuales.

En 2019 Andrango Calvache, Mónica Alexandra y Recalde Calvache, Pablo Marcel de la Universidad Tecnológica Israel, en el artículo de investigación “USO DE RECONOCIMIENTO FACIAL DE EMOCIONES BASADO EN TÉCNICAS DE DEEP LEARNING PARA EL MEJORAMIENTO DE LA EDUCACIÓN” desarrollado en la ciudad de Quito, se abordó el estudio de la detección de emociones en. La metodología se basó en la utilización de herramientas tecnológicas específicas, como Python con OpenCV, en combinación con un microcontrolador Raspberry Pi, con el propósito de obtener y procesar fotografías. Estas imágenes se sometieron a un análisis que permitió generar un histograma que representara las emociones manifestadas por un grupo de 234 estudiantes que cursaban el primer año de bachillerato. Los resultados obtenidos revelaron una predominancia significativa de la emoción de tristeza, que se manifestó en un 47% en comparación con las emociones neutras y positivas. Este hallazgo condujo a la consideración de cambios en las estrategias pedagógicas [5].

En 2020 Medina Guevara, María Gabriela y Amaya Preciado, Gonzalo Renato de la Universidad Jorge Tadeo Lozano, en la investigación “INTELIGENCIA EMOCIONAL ARTIFICIAL” se llevó a cabo una simulación que implicó la implementación de un mentor virtual, con el propósito de detectar las emociones de los trabajadores de un hospital. Esta detección se basó en la utilización de inteligencia artificial, particularmente en algoritmos diseñados para identificar estados emocionales a través de las respuestas proporcionadas por los trabajadores. Los datos resultantes del algoritmo de inteligencia artificial, que se basan en las respuestas proporcionadas por cada empleado del hospital al evaluar su estado emocional y anímico, mostraron resultados altamente significativos. Se reveló que el 65% de los

empleados no se encontraba en un estado emocional eficiente, experimentando cambios emocionales notables. Por otro lado, el 20% de los empleados mantenía emociones positivas, como la alegría, y expresaba satisfacción en su trabajo, mientras que el 15% restante experimentaba un estado emocional negativo, caracterizado por la tristeza, lo que sugiere que su motivación para trabajar estaba impulsada principalmente por necesidad. Estos resultados proporcionaron información valiosa que se utilizó para influir de manera positiva en la rutina diaria de los enfermeros, con el objetivo de promover un estado mental óptimo durante su desempeño laboral [6].

En el 2020 Pinzón Naranjo, Juan David y Sandoval, Nydia Susana de la Universidad de Pamplona, en la investigación “DESARROLLO DE SISTEMA PARA LA DETECCIÓN DE EMOCIONES UTILIZANDO VISIÓN ARTIFICIAL E INTELIGENCIA ARTIFICIAL APLICADO A ENTORNOS EDUCATIVOS” se presenta un sistema mediante el empleo de técnicas de inteligencia artificial, más específicamente, el aprendizaje automático a través de machine learning, utilizando la biblioteca OpenCV de Python. Este sistema se ha capacitado utilizando un conjunto de imágenes que abarcan un rango de 400 a 700 imágenes por cada emoción propuesta. Estas imágenes se han empleado para construir una base de datos integral que contribuye a una descripción más precisa de las emociones faciales. En una fase subsiguiente del proyecto, se ha utilizado la herramienta Tkinter para diseñar la interfaz destinada a la detección de emociones, lo que ha permitido obtener un sistema funcional y operativo. El aporte destacado de este proyecto radica en la identificación de un conjunto de 7 emociones, un aumento respecto a las 6 originalmente previstas. Este incremento se atribuye a la inclusión de una categoría adicional, denominada "neutro", que se manifiesta en situaciones donde el rostro presenta cambios faciales mínimos o poco perceptibles [7].

En 2022 Briones Pérez, Maraya Nataly y Moreno Armendáriz, Marco Antonio del Instituto Politécnico Nacional, en la investigación “DETECCIÓN DE EMOCIONES USANDO INTELIGENCIA ARTIFICIAL PARA LA RECOMENDACIÓN PERSONALIZADA DE BEBIDAS” en la ciudad de México, se llevó a cabo una investigación orientada al desarrollo de modelos destinados a la identificación de las emociones de los individuos, con el propósito de ofrecer recomendaciones de bebidas de bubble tea. Para llevar a cabo esta tarea, se empleó inteligencia artificial, haciendo



uso del lenguaje de programación Python, con especial atención en el entrenamiento de redes neuronales convolucionales. Los resultados resaltaron la relevancia del establecimiento de un entendimiento mutuo entre los consumidores y los agentes de servicios en el marco del marketing relacional. Se llevó a cabo la evaluación del estado emocional de 200 individuos con el propósito de facilitar respuestas empáticas y personalizadas a los usuarios, con el objetivo de generar una experiencia agradable al consumir la bebida. No obstante, se observó que solo el 45% de los individuos experimentó una mejora en su percepción y satisfacción con la bebida, mientras que los demás mantuvieron una opinión inalterada sobre la misma, aunque la consumieron [8].

En 2023 Flores Masías, Edward José y Livia Segovia, José Héctor, de la Universidad Nacional Federico Villareal, en la investigación “ANÁLISIS DE SENTIMIENTOS CON INTELIGENCIA ARTIFICIAL PARA MEJORAR EL PROCESO ENSEÑANZA-APRENDIZAJE EN EL AULA VIRTUAL” se centraron en el análisis de las emociones experimentadas por los estudiantes durante las clases en línea, utilizando una solución de software de bajo costo desarrollada en Python, que integra inteligencia artificial, en particular, el entrenamiento de una red neuronal convolucional para la detección de emociones. La implementación de esta solución no permitió la detección en tiempo real, en su lugar, capturó imágenes a intervalos específicos durante las clases en línea. Estas imágenes se sometieron a un análisis dentro de la aplicación, que posteriormente transmitió únicamente el estado emocional detectado al docente. Los datos obtenidos revelan que más del 85% de los estudiantes experimentan variaciones emocionales durante las clases en línea, mayoritariamente debido a dificultades en su nivel de concentración y distracciones presentes en su entorno de estudio. Esta información posee un valor sustancial al ofrecer oportunidades para la mejora de las estrategias pedagógicas aplicadas en el entorno virtual. Los docentes pudieron aprovechar estos datos para tomar decisiones en tiempo real y adaptar la dinámica de enseñanza, con el propósito de potenciar un proceso de aprendizaje más efectivo. Este enfoque concuerda con la visión de una formación integral del estudiante, representando un desafío esencial para el futuro de la educación en línea [9].

## 1.3 Fundamentación teórica

### 1.3.1 Emociones

Las emociones se definen como respuestas psicofisiológicas que representan estrategias adaptativas del individuo cuando este percibe un objeto, persona, lugar, evento significativo o recuerdo. Desde una perspectiva psicológica, las emociones influyen en la atención, potencian ciertas pautas conductuales y activan conexiones relevantes en la memoria. Además, desencadenan una serie de respuestas en diversos sistemas biológicos, como expresiones faciales, actividad muscular, vocalización, el sistema nervioso autónomo y el sistema endócrino, con el propósito de establecer un entorno interno óptimo para una conducta efectiva [10].



Figura 1. Las emociones en las personas [10].

#### *a. Tipos de emociones*

Las teorías sobre emociones fundamentales sugieren la presencia de emociones elementales o discretas, las cuales tienen la capacidad de cambiar en intensidad y combinarse para formar emociones más elaboradas y matizadas. En este contexto, Paul Ekman y su grupo, han identificado patrones para seis emociones fundamentales que aparentan ser biológicamente primarias y universales en todas las culturas [11].

En la Figura 2 se muestran las emociones universales propuesta por Ekman.



Figura 2. Rueda de Emociones de Ekman [11].

- Emociones principales.** La compilación de emociones elementales conocida como "Las Seis Grandes Emociones" ha sido ampliamente aceptada. Estas emociones se consideran fundamentales tanto desde una perspectiva psicológica como biológica, ya que no se descomponen en otras emociones preexistentes y se cree que son innatas. La propuesta de Ekman y su equipo ha dejado una marca significativa en la investigación subsiguiente, la cual ha buscado identificar patrones emocionales con respuestas distintivas en el sistema nervioso autónomo, respaldando así la concepción de emociones básicas. En la Tabla 1 se muestran las emociones principales universales [11].

Tabla 1. Emociones principales [11].

Emoción	Descripción
<b>Felicidad</b>	Proporciona una sensación de bienestar y satisfacción, además de fomentar la conexión social.
<b>Asco</b>	Es un fuerte desagrado experimentado hacia algo que resulta repulsivo.
<b>Enojo</b>	Esta emoción se caracteriza por la irritabilidad o el enojo y desencadena una función adaptativa de autodefensa.
<b>Miedo</b>	El recelo o aprensión frente a una amenaza o posible peligro, incluso si es imaginario, se traduce en una emoción que genera angustia, inseguridad e incertidumbre.
<b>Sorpresa</b>	Se trata de una emoción que surge ante un acontecimiento inesperado.
<b>Tristeza</b>	Emoción que te causa infelicidad y desdicha, especialmente cuando experimentas la pérdida de algo o alguien que tiene un gran significado para ti.

Cada emoción descrita en la Tabla 1 cumple una función específica y se manifiesta de manera única, lo cual puede ser de considerable utilidad para elevar nuestra

inteligencia emocional y nuestra destreza en el manejo efectivo de nuestras reacciones emocionales [11].

- **Emociones secundarias.** Surgen como una evolución posterior a las emociones fundamentales, producto de nuestra adaptación social y del desarrollo de nuestras facultades cognitivas como se muestra en la Tabla 2 [11].

Tabla 2. Emociones secundarias [11].

<b>Emoción</b>	<b>Descripción</b>
<b>Culpa</b>	Aparece cuando te das cuenta de que has transgredido tus propios principios personales y los valores comunes de la sociedad a la que perteneces.
<b>Celos</b>	Estas emociones se hacen evidentes cuando notas la eventualidad de perder algo significativo o que valoras como propio.
<b>Vergüenza</b>	Surge cuando anticipas que las personas cercanas a ti, o aquellas que podrían estarlo, presenciarán una acción o evento que perjudicará tu propia reputación o la de alguien más.
<b>Orgullo</b>	Se trata de una reacción adaptativa que se manifiesta cuando te sientes sumamente satisfecho con tu propio ser o con tus éxitos.
<b>Placer</b>	Es una sensación positiva que se experimenta al lograr satisfacción.

### 1.3.2 Características del rostro

Al detectar un rostro humano, se sigue un proceso esencial que implica el reconocimiento de diversos elementos presentes en el rostro durante nuestras actividades cotidianas. Mediante diversos métodos, se logra reconocer los atributos y características específicas que cada rostro posee, a continuación en la Tabla 3 se presentan las características del rostro [12].

Tabla 3. Atributos principales de la detección del rostro [12].

<b>Atributos</b>	<b>Descripción</b>
<b>Expresiones faciales</b>	Las expresiones faciales se relacionan con las modificaciones que ocurren en nuestro rostro como respuesta a estados emocionales internos. En este sentido, examinar las expresiones faciales implica identificar de manera efectiva estas manifestaciones en el rostro.
<b>Puntos de referencia</b>	Son áreas particulares en el rostro que se emplean como puntos de referencia cruciales para el análisis y la comparación en el reconocimiento facial.
<b>Proporciones faciales</b>	Hacen referencia a las proporciones y dimensiones entre las diversas partes del rostro, abarcando aspectos como la distancia entre los ojos y la relación entre la longitud de la nariz y la boca.
<b>Perfiles faciales</b>	Estas se vinculan con la visualización del rostro de una persona, indicando la observación de la cara desde un ángulo lateral en lugar de una vista frontal, lo cual revela la configuración de la nariz, la barbilla, la frente y otros rasgos que pueden resultar útiles.

### 1.3.3 Indicadores faciales para la detección de emociones

Los indicadores faciales de la Figura 3 que se utilizan para la detección de emociones son una serie de señales y expresiones visibles en el rostro humano que revelan el estado emocional de una persona en un momento dado. Estos indicadores comprenden expresiones faciales específicas, gestos y modificaciones en los músculos faciales. Su relevancia se destaca en disciplinas como la psicología, la inteligencia artificial y la investigación emocional, ya que posibilitan inferir y comprender de manera más precisa los estados emocionales de las personas. Esto resulta valioso en diversas aplicaciones, desde la atención médica hasta la interacción humano-máquina [13].



Figura 3. Indicadores faciales [13].

#### *a. Movimiento de músculos faciales*

El rostro humano común cuenta con 43 músculos distintos que pueden activarse alrededor de los ojos, la nariz, la boca, la mandíbula, el mentón y la frente, posibilitando la creación de miles de expresiones diversas. Los músculos que se ocupan en las expresiones facial forman un conjunto de alrededor de 20 músculos planos ubicados subcutáneamente debajo de la piel de la cara y el cuero cabelludo. La mayoría de estos músculos se originan en estructuras óseas o fibrosas del cráneo y se conectan a la piel circundante [14].

A diferencia de los músculos esqueléticos tradicionales, en la anatomía facial, los músculos de la Tabla 4 no están cubiertos por fascia, excepto el músculo buccinador. Estos músculos se distribuyen alrededor de las aperturas naturales del rostro, como la boca, los ojos, la nariz y los oídos, o se extienden a lo largo del cráneo y el cuello [14].

Tabla 4. Movimientos principales de los músculos en las expresiones faciales [14].

Movimientos	Características
<b>Elevación de las cejas</b>	Se refiere al movimiento hacia arriba de las cejas, que puede denotar sorpresa, atención o interés. También puede utilizarse para expresar escepticismo o incredulidad en ciertas circunstancias.
<b>Apertura de los ojos</b>	Se relaciona con la expansión de los párpados superiores e inferiores. Un aumento en la apertura de los ojos puede indicar sorpresa, alerta o asombro, mientras que una disminución en la apertura puede sugerir cansancio, aburrimiento o relajación.
<b>Arrugas en la nariz</b>	Se forman al fruncir la nariz o arrugarla, y suelen estar asociadas con la expresión de desprecio o disgusto. Este movimiento puede ser una respuesta a olores desagradables o situaciones incómodas.
<b>Movimiento de la boca</b>	Abarca una variedad de expresiones faciales, incluyendo sonrisas, muecas, fruncimientos y otras formas de manipulación de los labios y la boca. Cada movimiento de la boca puede transmitir una emoción diferente, desde alegría y felicidad hasta tristeza y enojo.

Para llevar a cabo la detección de indicadores faciales en el rostro humano, es necesario realizar una serie de procesos que contribuirán a identificar los rasgos más característicos, especialmente en relación con las emociones específicas. En este contexto, una de las técnicas destacadas en el proceso de detección es el empleo del procesamiento digital de señales e imágenes [14].

### 1.3.4 Procesamiento digital de Señales

La base del Procesamiento de Señales Digitales, conocido por sus siglas en inglés Digital Signal Processing (DSP), reside en la manipulación matemática de una señal de información con el propósito de reducir el ruido, mejorarla en aspectos particulares y combinar sus valores para generar nuevos. Su característica distintiva principal se encuentra en la representación de las señales en el dominio de tiempo discreto, dominio de frecuencia discreta u otro dominio discreto mediante el uso de valores numéricos o símbolos [15].

El DSP se encarga de convertir las señales analógicas en datos digitales y emplea algoritmos y operaciones matemáticas para analizar, ajustar, modificar y adaptar las señales según las necesidades específicas. Este procedimiento se lleva a cabo mediante un sistema que incorpora un procesador o microprocesador, el cual recopila información de una fuente y la procesa de manera rápida y en tiempo real después de la conversión a formato digital. Dada su considerable versatilidad, el procesamiento digital de señales tiene la capacidad de ejecutar tareas complejas en tiempo real,

convirtiéndose en una herramienta esencial en campos como las comunicaciones, el control industrial y el procesamiento de audio y video, entre otros. A continuación en la Figura 4 se presenta un sistema digital que procesa datos [16].

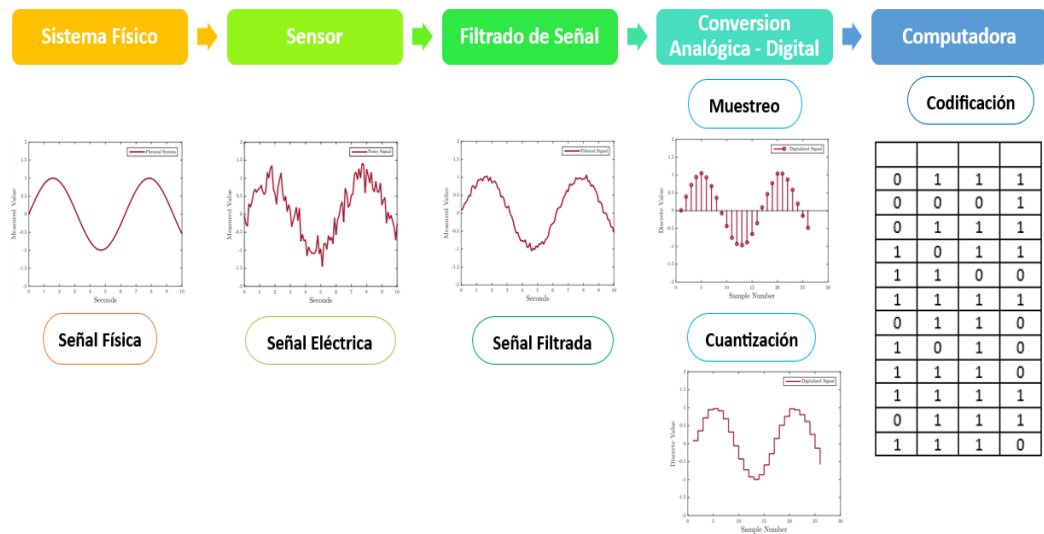


Figura 4. Sistema Digital de Adquisición de Datos [17].

#### a. *Procesamiento digital de imágenes*

El procesamiento digital de imágenes abarca el conjunto de técnicas y procedimientos utilizados para revelar o resaltar la información contenida en una imagen. En la actualidad, este campo de estudio representa una especialización destacada en el ámbito informático y está estrechamente vinculado al procesamiento digital de señales. La evolución de las técnicas de procesamiento digital de imágenes se origina en dos áreas principales de aplicación: la mejora de la información visual para la interpretación humana y el tratamiento de datos de la imagen para la percepción de máquinas autónomas, que incluye fases de transmisión y/o almacenamiento de dichos datos, en la Figura 5 se presenta un ejemplo del procesamiento digital de imágenes [18].



Figura 5. Procesamiento digital de imágenes [19].

- **Características del procesamiento digital de imágenes.** En el ámbito del procesamiento de imágenes, se utilizan diversas técnicas que comprenden operaciones basadas en una imagen, lo que conduce a la generación de una nueva imagen. El valor de cada píxel en la imagen resultante puede depender de su valor en la imagen original, de los valores de los píxeles circundantes o incluso de los valores de todos los puntos en la imagen de origen [20].

Estas técnicas buscan modificar una imagen de manera que la nueva imagen sea más adecuada para una aplicación específica en comparación con la imagen original. La palabra "específica" es de suma importancia, ya que señala que el valor de la imagen resultante se determina según el problema abordado, la siguiente Tabla 5 proporciona las principales características del procesamiento de imágenes [20].

Tabla 5. Características del procesamiento digital de imágenes [20].

<b>Características</b>	<b>Descripción</b>
<b>Digitalización de Imágenes</b>	Involucra la transformación de imágenes analógicas a formatos digitales mediante el uso de técnicas de muestreo y cuantización. La estructura básica implica la creación de conjuntos de elementos (vectores, matrices, tensores) organizados de manera ordenada.
<b>Representación en Píxeles</b>	Las imágenes se expresan a través de una matriz de píxeles, donde cada píxel contiene datos relativos al color y la intensidad lumínica en una posición particular.
<b>Histograma de una imagen</b>	La representación numérica de una imagen mediante una disposición de N filas por M columnas resulta en una matriz M×N. El histograma de una imagen, también conocido como "mejora de imagen" o "caracterización de imagen" y referido como h(i), es un vector que indica la cantidad de píxeles en la imagen con un valor particular en sus componentes.
<b>Resolución de imágenes</b>	La definición de la resolución de una imagen se basa en la cantidad de píxeles que la componen, y su unidad común de medida es "píxeles por pulgada" (ppi). En consecuencia, tanto la calidad de la representación como las dimensiones de la imagen están vinculadas a la resolución, la cual impacta en los requisitos de memoria para el archivo gráfico a generar.

- **Aplicaciones del procesamiento digital de imágenes.** El uso del procesamiento digital de imágenes ha evolucionado en contraste con sus características, adaptándose y mejorando su desempeño. A continuación, en la Tabla 6 se presentan algunas de las numerosas aplicaciones en la actualidad [21].



Tabla 6. Aplicaciones del procesamiento digital de imágenes [21].

<b>Aplicaciones</b>	<b>Descripción</b>
<b>Transformaciones Geométricas</b>	Simplifica las alteraciones geométricas, como rotación, escalado, traslación y deformación, con el fin de cambiar la estructura geométrica de la imagen.
<b>Segmentación</b>	Su aplicación consiste en la partición de la imagen en áreas o segmentos que comparten características similares, simplificando así el análisis y la identificación de objetos.
<b>Reconocimiento de Patrones</b>	Contribuye al reconocimiento de patrones y objetos presentes en una imagen, aspecto crucial en aplicaciones como la visión por computadora y la inteligencia artificial.
<b>Compresión de Imágenes</b>	Facilita la compresión de imágenes con el objetivo de disminuir el tamaño del archivo, simplificando así el almacenamiento y la transmisión.
<b>Restauración de Imágenes</b>	Se emplea para recuperar imágenes que han sufrido daños o degradación debido a factores como el ruido o la pérdida de información durante la transmisión.

Las características mencionadas posicionan al procesamiento digital de imágenes como una herramienta esencial en una amplia variedad de campos, que incluyen, entre otros, la medicina, la investigación científica, la visión por computadora y el ámbito del entretenimiento.

- **Procesamiento digital de imágenes con inteligencia artificial.** En las últimas décadas, la incorporación de la inteligencia artificial (IA) ha ocasionado una transformación notable en el campo del procesamiento digital de imágenes. La fusión de capacidades avanzadas de cómputo y algoritmos de aprendizaje automático ha marcado el comienzo de una nueva era en la manipulación de imágenes, posibilitando que los sistemas no solo manejen datos visuales, sino que también aprendan y se ajusten de forma autónoma [22].

La inteligencia artificial ha proporcionado una capacidad sin precedentes para entender y extraer características complejas de las imágenes, dando lugar a avances notables en áreas como el reconocimiento de objetos, la segmentación semántica, la mejora automática de imágenes y la generación de contenido visual. Este enfoque ha transformado la manera en que abordamos la interpretación y manipulación de imágenes digitales, abriendo la puerta a diversas aplicaciones innovadoras en campos que incluyen la medicina, la investigación científica, la realidad virtual y el entretenimiento digital. La aplicación computacional de técnicas y métodos de inteligencia artificial en imágenes implica seguir una secuencia de procedimientos metodológicamente establecidos, como se describe en la Tabla 7 [22]:

Tabla 7. Procedimientos del procesamiento de imágenes con IA [22].

Procedimientos	Descripción
<b>Obtención de descriptores</b>	Es posible obtener descriptores para cada elemento que se busca identificar en las imágenes, pueden evaluarse mediante la clasificación de características locales, tales como el color, la reflectancia en cada banda, la dimensión fractal, entre otros. Esta clasificación puede llevarse a cabo en espacios específicos, mediante componentes principales u otras transformadas.
<b>Segmentación</b>	Este procedimiento consiste en generar una región en la imagen procesada que contenga descriptores clasificados de manera uniforme y posea una geometría lo más sencilla posible para su representación. Las metodologías adecuadas para llevar a cabo esta tarea se basan en la identificación de discontinuidades y su conexión mediante algoritmos de seguimiento inteligente.
<b>Reconocimiento</b>	Es crucial conectar las diversas áreas segmentadas con objetos o características identificadas previamente durante una etapa de entrenamiento y almacenadas en una base de conocimientos. Esto requiere el uso de procesamiento inteligente de alto nivel, que abarca inferencia lógica, marcos conceptuales y sistemas expertos.
<b>Extracción y medición de características</b>	Después de identificar los distintos elementos que se buscan clasificar en las imágenes, se lleva a cabo la medición de sus características esenciales para la evaluación de parámetros en los modelos. Este procedimiento implica la aplicación de técnicas de estimación estadística con diversas distribuciones.

### 1.3.5 Inteligencia Artificial

La inteligencia artificial se establece como la base para simular los procesos de inteligencia humana mediante la creación y aplicación de algoritmos desarrollados en un entorno computacional dinámico. En términos más sencillos, la IA tiene como objetivo que las computadoras puedan pensar y actuar de manera similar a los seres humanos. Para lograr esto, se necesitan sistemas computacionales, la gestión eficiente de datos y algoritmos avanzados de IA. La búsqueda de una similitud significativa con el comportamiento humano implica la necesidad de mayores cantidades de datos y capacidad de procesamiento [23].

#### *a. Visión Artificial*

La visión artificial, un subcampo de la inteligencia artificial (IA), enseña a las computadoras y sistemas a extraer información significativa de imágenes digitales, videos y otras entradas visuales. Esto permite tomar medidas o dar recomendaciones basadas en la información visual obtenida. Mientras que la IA habilita a las computadoras para pensar, la visión artificial les confiere la capacidad de ver, observar y comprender, la Figura 6 representa un diagrama básico acerca de la visión artificial [24].

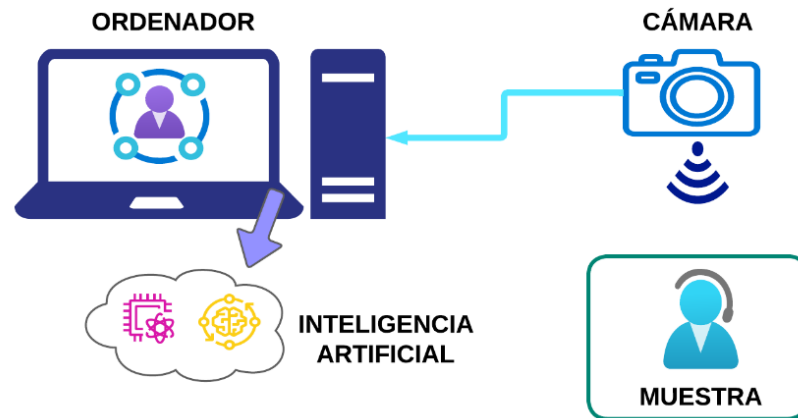


Figura 6. Diagrama básico de la visión artificial [24].

El funcionamiento de la visión artificial comparte similitudes con la visión humana, aunque los humanos tienen una ventaja inicial. La vista humana se beneficia de la experiencia acumulada a lo largo de toda una vida para aprender a distinguir objetos, estimar distancias, detectar movimientos e identificar anomalías en una imagen [24].

- **Reconocimiento Facial.** El reconocimiento facial representa una aplicación específica de la visión artificial que se enfoca en la identificación y verificación de individuos mediante el análisis de sus características faciales distintivas. Esta tecnología utiliza algoritmos y modelos de inteligencia artificial para adquirir, analizar y comparar atributos faciales, como la forma de los ojos, la nariz, la boca y la disposición de puntos de referencia, con el propósito de autenticar o identificar a una persona. El reconocimiento facial se aplica en ámbitos como la seguridad biométrica, sistemas de acceso, videovigilancia y diversas áreas más [25].

Dada su integración con la inteligencia artificial, los sistemas de reconocimiento facial ofrecen niveles significativos de seguridad y fiabilidad, convirtiéndolos en alternativas preferentes para una variedad de aplicaciones, que abarcan desde la autenticación en dispositivos móviles hasta la verificación de identidad en plataformas en línea y eventos. Además de optimizar la toma de decisiones mediante la reducción de la carga de trabajo manual, estos sistemas posibilitan que los profesionales dirijan su atención hacia actividades de índole más estratégica, en la Figura 7 se describe el diagrama básico del reconocimiento facial [25].

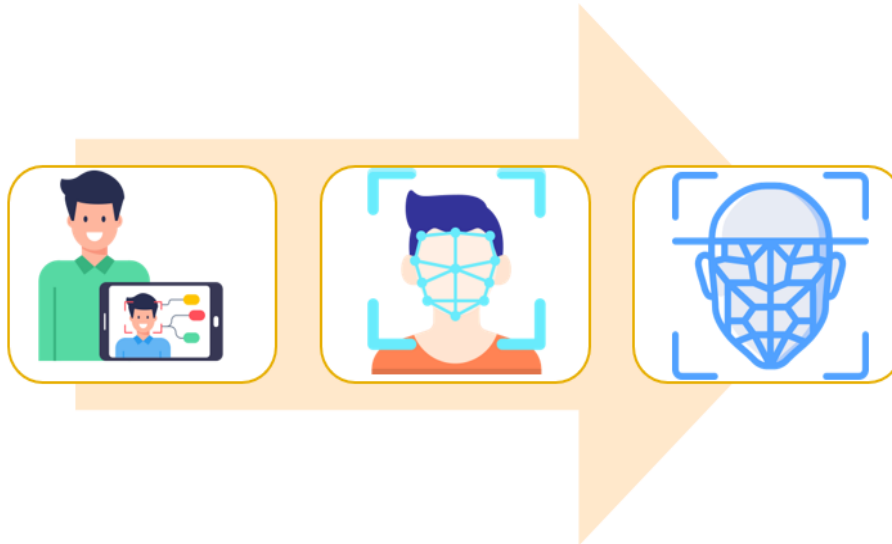


Figura 7. Diagrama básico del Reconocimiento facial [25].

Existen dos enfoques para llevar a cabo el reconocimiento facial, los cuales dependen de la tecnología facial disponible [25].

- El primer método, que era ampliamente utilizado hasta hace poco años, se basa en un detector facial a través de fotografías o frames. Esta metodología fue prevalente antes de que los avances en inteligencia artificial, visión artificial y aprendizaje automático comenzaran a impactar significativamente en la tecnología facial [25].
- En la actualidad, el método más seguro y confiable de reconocimiento facial utiliza tecnología basada en vídeo. Los sistemas de reconocimiento facial basados en video generalmente se componen de tres módulos: uno dedicado a la detección facial, otro encargado del seguimiento, y un tercero destinado al reconocimiento [25].

En la mayoría de estos sistemas, se elige un conjunto de cuadros de calidad en los cuales se aplican técnicas de reconocimiento facial para la identificación de la persona. El proceso de reconocimiento facial se compone de tres fases que se presentan en la Figura 8 para cumplir con su objetivo [26].

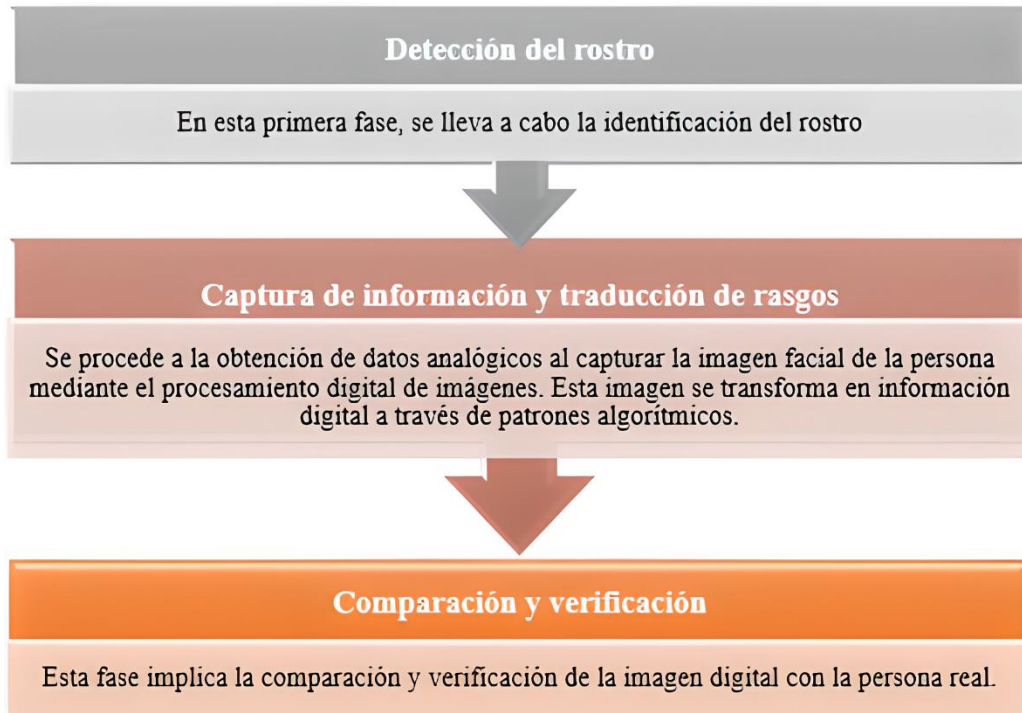


Figura 8. Proceso de reconocimiento facial [25].

La eficacia de la tecnología de reconocimiento facial se ha destacado debido a su robustez, fiabilidad y baja intrusión. Además, el proceso de adquisición y comparación de información biométrica es simplificado, ya que un solo retrato de la persona es suficiente para lograr los objetivos, como la identificación, comparación o verificación [26].

- **Componentes faciales de interés.** Hacen referencia a los órganos y elementos constitutivos de una cara humana típica, como cejas, ojos, boca, nariz, entre otros. No existe una elección única y universal de estos componentes faciales, ya que diferentes métodos optan por distintos elementos mientras excluyen algunos. La cantidad de componentes a considerar varía según el nivel de detalle necesario para una aplicación específica. Por ejemplo, para una interfaz perceptual, puede ser suficiente conocer la posición media de ambos ojos, pero en un sistema de análisis de la dirección de la mirada, es esencial distinguir entre el globo ocular, el párpado, el iris y la pupila. Entre los elementos faciales más relevantes en la literatura, se destacan en la Tabla 8 [27]:

Tabla 8. Componentes faciales importantes [27].

Componentes	Descripción
<b>Ojos</b>	La ubicación se describe comúnmente en términos de la posición media, aunque definirlo con precisión puede presentar un desafío. También surge la cuestión de cómo definir la posición cuando un ojo está cerrado. La referencia a un "ojo izquierdo" o "ojo derecho" se basa en su posición en la imagen, lo cual puede ser contrario a la perspectiva de la persona retratada. Otro aspecto relevante es la orientación de la cara en la imagen, que se mide en función del ángulo formado por una línea que pasa a través de ambos ojos en relación con el eje horizontal.
<b>Boca</b>	La descripción de la ubicación de la boca se aborda de diversas maneras, siendo la más básica la posición media de la boca o los labios. Sin embargo, esta elección puede generar ambigüedades, especialmente cuando la boca está abierta. Algunos enfoques también tienen en cuenta la extensión horizontal y vertical de la boca, la cual está vinculada a su grado de apertura.
<b>Nariz</b>	La identificación de la ubicación de la nariz suele ser de menor relevancia debido a su movilidad limitada y a que su expresión facial está comúnmente asociada con la boca. Cuando se considera, por lo general se busca la posición de la punta de la nariz o los orificios nasales.

- **Puntos de referencia.** Son ubicaciones específicas que se observan en la Figura 9 en el rostro que se utilizan como puntos clave para el análisis y la comparación en el reconocimiento facial [27].

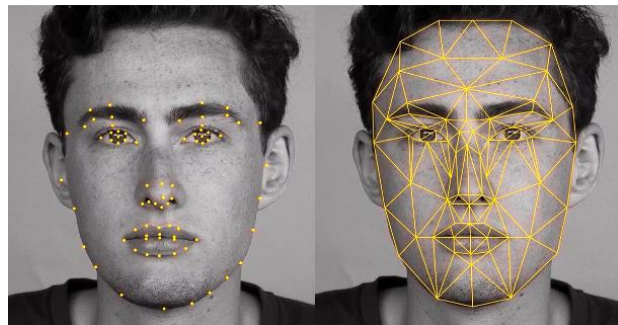


Figura 9. Puntos de referencia del rostro [28].

- **Reconocimiento de patrones de las emociones.** El procedimiento implica la extracción y evaluación de datos significativos de una representación facial, transformando estas características en información digital que puede ser comparada y utilizada para identificar a un individuo específico. Los algoritmos de inteligencia artificial son entrenados para reconocer patrones mediante el uso extensivo de conjuntos de datos amplios que contienen ejemplos de rostros. Estos modelos se aplican posteriormente en tareas como la identificación de una persona en una imagen, la verificación de su identidad o la comparación con otras imágenes para determinar similitudes, la Tabla 9 proporciona patrones faciales de las emociones [29].

Tabla 9. Patrones faciales de las emociones [7] [8] [29].

Emoción	Definición	Patrones faciales
<b>Ira</b>	La sensación de hostilidad hacia una persona o cosa suele surgir después de experimentar una percepción de perjuicio u ofensa.	Disminuir el espacio entre las cejas, contraer los labios, mantener la mirada de forma intensa y cerrar ocasionalmente los párpados inferiores, a veces acompañado por el desplazamiento hacia adelante de la mandíbula.
<b>Tristeza</b>	Emoción de desdicha o melancolía.	Los párpados bajos, las comisuras de la boca hacia abajo, los labios fruncidos y la mirada dirigida hacia abajo.
<b>Sorpresa</b>	Una percepción desconcertante o de asombro ante un suceso imprevisto.	Levantar las cejas de manera notable (lo cual podría ocasionar arrugas en la frente), abrir los ojos de manera extensa y permitir que la mandíbula se relaje en una posición inferior.
<b>Miedo</b>	Una impresión de inquietud originada por la percepción de riesgo amenaza o la posibilidad de experimentar malestar.	Levantar las cejas o contraerlas, apretar los párpados inferiores, extender los labios horizontalmente y mantener la boca ligeramente abierta.
<b>Alegría</b>	Una agradable sensación de satisfacción y bienestar.	Al exhibir una sonrisa, elevando las esquinas de los labios y activando los músculos orbitales alrededor de los ojos.

### 1.3.6 Machine Learning

Es la capacidad de un sistema de inteligencia artificial para adquirir conocimiento de manera autónoma a través de un ciclo de aprendizaje que implica la utilización de datos, procesos de entrenamiento y la obtención de resultados. Hay varias variantes, dependiendo de si el aprendizaje requiere supervisión humana o se permite que la IA adquiera conocimientos de forma autodidacta, de acuerdo con reglas predefinidas. Este enfoque se utiliza frecuentemente en aplicaciones como asistentes virtuales y chatbots, entre otras, en la Figura 10 se detallan las características del machine learning [30].



Figura 10. Machine Learning [31].

La concepción de aplicaciones de machine learning, a pesar de compartir similitudes con el paradigma convencional de ingeniería, presenta una distinción importante: abordar los datos como componente fundamental. La efectividad de un proyecto recae, en gran medida, en la calidad y gestión de los datos adquiridos.

**a. Tipos de machine Learning**

En la tabla siguiente se exhiben distintos tipos de aprendizaje automático (machine learning) que desempeñan un papel fundamental en el ámbito de la inteligencia artificial. Cada enfoque presenta características particulares y aplicaciones específicas. La clasificación comprende desde el aprendizaje supervisado, en el cual el modelo se entrena con ejemplos etiquetados, hasta el aprendizaje no supervisado, que investiga patrones sin etiquetas previas. Asimismo, se integra el aprendizaje por refuerzo, que se fundamenta en la interacción del modelo con su entorno para mejorar su rendimiento. A continuación, en la Tabla 10 se detallan y explican estos enfoques, brindando una perspectiva exhaustiva de las diversas metodologías en el campo del aprendizaje automático [32].

Tabla 10. Tipos de aprendizaje de automático [33].

Tipos de Machine Learning	Descripción
<b>Aprendizaje supervisado</b>	Modalidad de aprendizaje en la cual se dispone de la totalidad de la información correspondiente a los datos de entrenamiento, abarcando tanto las entradas como las salidas asociadas. Esta forma de aprendizaje destaca por proporcionar resultados óptimos, dado que posee la máxima cantidad de información disponible.
<b>Aprendizaje no supervisado</b>	Modalidad de aprendizaje en la cual se cuentan exclusivamente con los datos de entrada, siendo el propósito principal obtener información acerca de la estructura del dominio de salida.
<b>Aprendizaje semi-supervisado</b>	Se trata de una forma de aprendizaje que presenta características híbridas entre los enfoques supervisado y no supervisado.
<b>Aprendizaje adaptativo</b>	Modalidad de aprendizaje que inicia con un modelo preexistente, cuyos parámetros son ajustados o adaptados mediante la incorporación de nuevos datos durante el proceso de entrenamiento.
<b>Aprendizaje por refuerzo</b>	Modalidad de aprendizaje que se sitúa en una intersección entre el aprendizaje en línea y el aprendizaje semisupervisado, caracterizado por una supervisión incompleta, usualmente expresada en términos binarios como {verdadero, falso}.



- **Características de los aprendizajes automáticos.** Los desafíos en el ámbito del aprendizaje automático se centran en la anticipación de las características o el comportamiento de un conjunto de datos específico. Estos problemas se pueden clasificar en dos categorías principales como se indica en la Figura 11 [33]:



Figura 11. Aprendizaje supervisado vs no supervisado [33].

En la Tabla 11 se presenta los algoritmos juntamente con su categoría vinculada [33]:

Tabla 11. Algoritmos de aprendizaje automático [33].

Algoritmo	Categoría
<b>K-Means</b>	Clustering
<b>Gaussian Mixture</b>	Clustering
<b>Redes Neuronales Convolucionales (CNN)</b>	Clustering
<b>Mínimos cuadrados ordinarios (OLS)</b>	Selección de características
<b>Naïve Bayes (NB)</b>	Clasificación
<b>K-Vecinos más cercanos (k-NN)</b>	Regresión, Clasificación
<b>Máquinas de vectores de soporte (SVM)</b>	Regresión, Clasificación
<b>Árboles de decisión (DT)</b>	Regresión, Clasificación
<b>Bosques Aleatorios (RF)</b>	Regresión, Clasificación
<b>Redes Neuronales Recurrentes (RNN)</b>	Regresión, Clasificación

### 1.3.7 Redes Neuronales Artificiales

En el ámbito biológico, las neuronas biológicas exhiben elementos esenciales como dendritas, el soma y el axón. Las dendritas cumplen la función de detectar e recibir impulsos nerviosos provenientes de otras neuronas. Estos impulsos se procesan en el soma y se transmiten a través del axón, generando impulsos nerviosos dirigidos hacia neuronas adyacentes. Las redes neuronales artificiales, inspiradas en el funcionamiento de las redes biológicas, constituyen conjuntos de elementos simples (generalmente adaptables) interconectados masivamente en paralelo, con una

estructura jerárquica, buscando emular la interacción con objetos del mundo real de manera análoga al sistema nervioso biológico [34].

***a. Ventajas de una red Neuronal***

Debido a su estructura y principios fundamentales, las redes neuronales artificiales comparten numerosas similitudes con el cerebro. Entre estas similitudes se destacan la capacidad para aprender de la experiencia, la habilidad para generalizar de casos previos a situaciones novedosas, y la capacidad para abstraer características esenciales a partir de entradas que representan información no relevante. Algunas de las ventajas se incluyen en la Tabla 12 [34]:

Tabla 12. Ventajas de una Red Neuronal Artificial [34].

<b>Ventajas</b>	<b>Descripción</b>
<b>Aprendizaje Adaptativo</b>	Esta característica se refiere a la habilidad de la red neuronal para adquirir conocimientos y ejecutar tareas basadas en el entrenamiento o la experiencia inicial, mediante la adaptación de sus conexiones y parámetros en concordancia con la información proporcionada.
<b>Auto-organización</b>	La red neuronal demuestra la aptitud para desarrollar su propia estructura organizativa o representación de la información en un período de aprendizaje, lo que le concede la capacidad de ajustarse y organizar de manera autónoma con el objetivo de mejorar su eficacia en tareas particulares.
<b>Tolerancia a fallos</b>	Aunque la degradación parcial de la red puede llevar a una pérdida de su estructura, ciertas capacidades esenciales pueden mantenerse incluso después de sufrir daños considerables, fortaleciendo su robustez y capacidad de recuperación.
<b>Operación en tiempo real</b>	La capacidad de realizar cálculos neuronales de forma concurrente posibilita la operación en tiempo real, y se construyen sistemas con hardware especializado para ejecutar estas operaciones de manera eficaz y veloz.
<b>Fácil inserción en la tecnología</b>	La disponibilidad de chips especializados para redes neuronales facilita su integración modular en sistemas existentes, mejorando el rendimiento en tareas específicas sin requerir modificaciones sustanciales en la infraestructura tecnológica preexistente.

***b. Arquitectura de una red Neuronal***

La arquitectura de una red neuronal artificial, modelada a partir de la complejidad del sistema nervioso biológico, desempeña un papel crucial en la comprensión de los procesos de aprendizaje y procesamiento de información. Al igual que el cerebro humano, una red neuronal artificial está compuesta por unidades fundamentales llamadas neuronas, que colaboran entre sí, para llevar a cabo funciones específicas

mediante la interconexión de capas, en la Figura 12 se detallan las capas de una red neuronal [34].

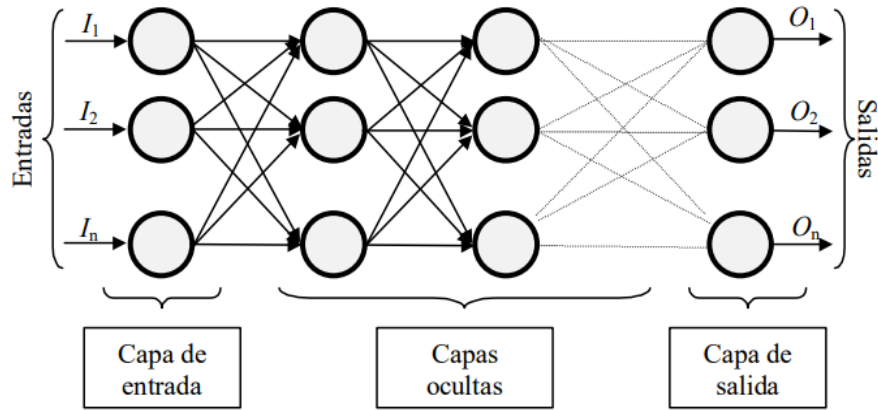


Figura 12. Arquitectura de una Red Neuronal [34].

Estas redes se organizan en capas, cada una con una función específica en el procesamiento de datos. La comprensión detallada de esta arquitectura resulta crucial para entender cómo estas redes tienen la capacidad de adquirir patrones complejos y llevar a cabo diversas tareas, desde el reconocimiento de patrones hasta la toma de decisiones. En la Figura 13, se analizó las tres capas esenciales que conforman la estructura de una red neuronal artificial [34].

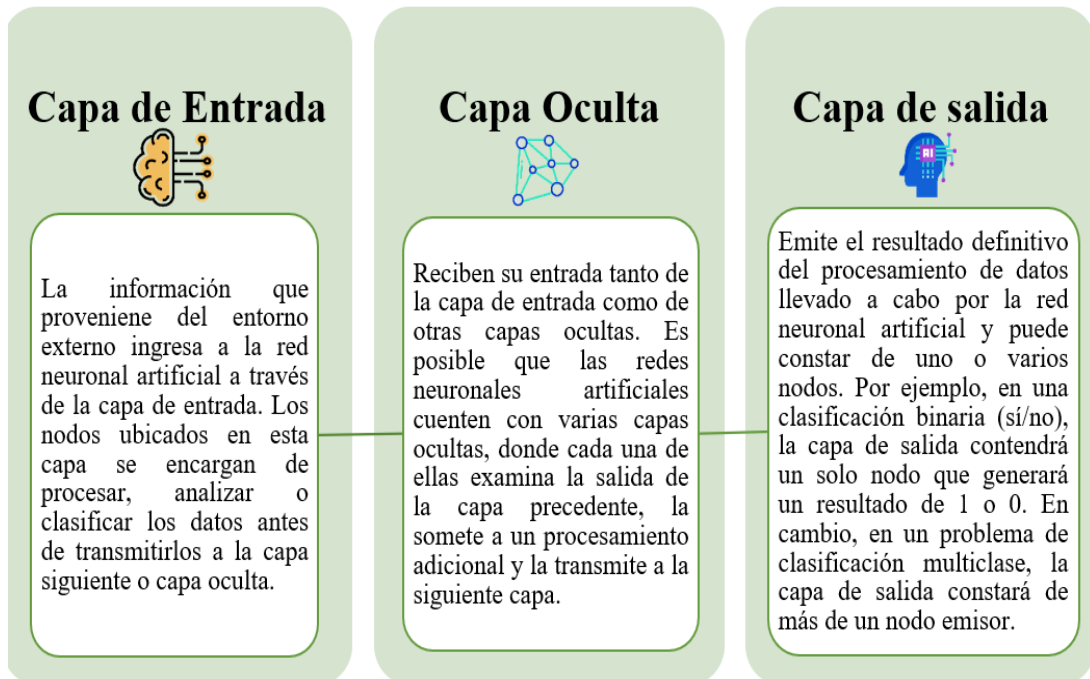


Figura 13. Características de las capas de una red neuronal artificial [34].

### 1.3.8 Redes Neuronales Convolucionales

Las Redes Neuronales Convolucionales representan una categoría de arquitecturas de redes neuronales profundas ideadas específicamente para el procesamiento de datos organizados en cuadrículas, como es el caso de las imágenes. Estas estructuras neuronales como se observa en la Figura 14 desempeñan un papel esencial en diversas actividades de visión por computadora, así como en el reconocimiento de patrones y el análisis de imágenes [33].

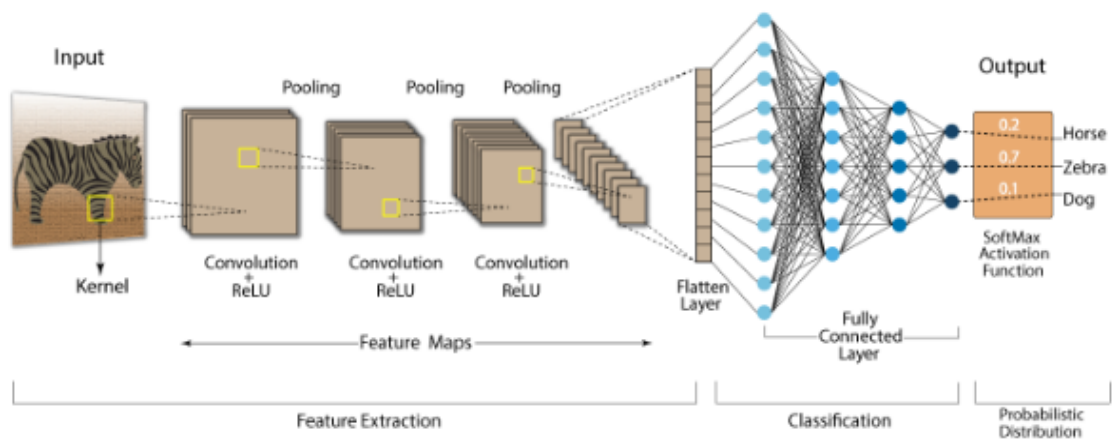


Figura 14. Diagrama de una Red Neuronal Convolutiva [35].

En el ámbito del aprendizaje profundo, se conoce como red neuronal convolutiva (CNN o ConvNet) ayudan en la arquitectura de pesos compartidos de los núcleos o kernels de filtros que se desplazan a lo largo de las características de entrada, generando respuestas equivalentes a desplazamientos conocidos como mapas de características. Estas redes encuentran aplicaciones en interfaces cerebro-computadora, reconocimiento facial, clasificación y segmentación de imágenes, análisis de imágenes, procesamiento del lenguaje natural, y análisis de series temporales financieras [33].

#### a. *Tensor Flow para redes neuronales convolucionales*

Una de las librerías más utilizadas en el desarrollo de redes neuronales es TensorFlow, diseñada específicamente para la ejecución eficiente de operaciones como la multiplicación de matrices que cuenta con una variedad de librerías para este tipo de

tareas, aprovechando de manera óptima el GPU o rendimiento de procesadores gráficos que son fabricados para estas tareas [33].

Su nombre se deriva de su estructura principal de datos: los "tensores". Este término tiene relación debido a que en una neurona se realizan operaciones sobre matrices de datos multidimensionales. Al basarse en grafos, estos representan operaciones matemáticas, mientras que sus conexiones son los datos multidimensionales, denominados tensores. Un tensor es una colección de datos primitivos, ya sea entero o flotantes arreglados en una matriz siendo el rango su dimensión abarcada. Estas estructuras de datos son fundamentales para facilitar el flujo de datos entre las capas de la red neuronal [33].

- **Arquitectura de tensor Flow.** Su arquitectura altamente adaptable posibilita la implementación de cómputo en una o varias unidades de procesamiento central (CPUs) o unidades de procesamiento gráfico (GPUs) en un servidor. El lenguaje de programación C++ es donde se realizan los cálculos de aprendizaje de la capa inicial. TensorFlow de NVIDIA hace uso de una biblioteca llamada CUDA, la capa intermedia, de Python, invoca un método TensorFlow que, a su vez, activa código C++. La capa final está compuesta por el API de Estimator, que incorpora diversas funciones y componentes para facilitar la construcción de modelos de aprendizaje automático, como se puede observar en la Figura 15 [33].

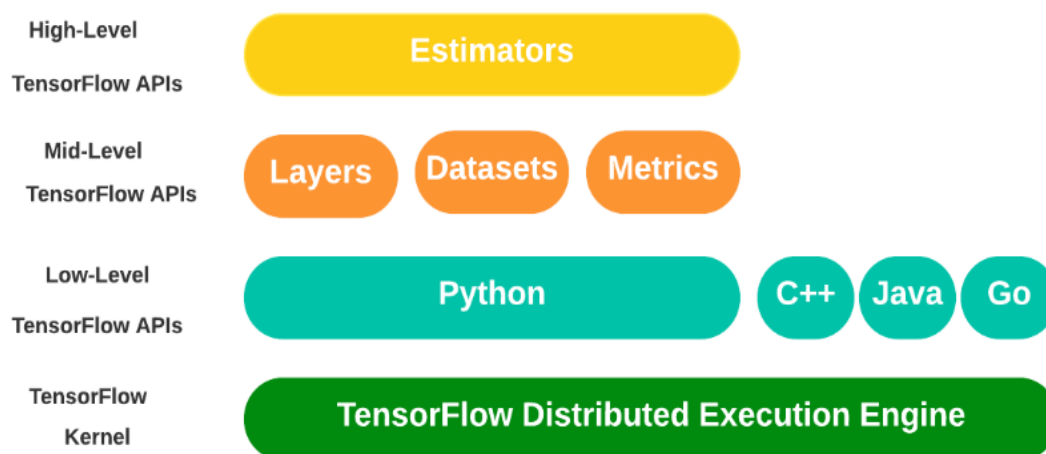


Figura 15. Arquitectura de TensorFlow [33].

- **Librería Keras en Deep learning.** En la implementación de una red neuronal en TensorFlow se empleará la biblioteca Keras para aprendizaje profundo. Keras es una interfaz de programación de aplicaciones (API) diseñada para redes neuronales, con la capacidad de ejecutarse en diversas plataformas de aprendizaje automático, siendo TensorFlow el enfoque primordial. Entre sus características destacadas se encuentran en la Figura 16 [33]:



Figura 16. Características de Keras en Deep Learning [33].

#### ***b. Validación de una red neuronal***

Después de completar el proceso de entrenamiento, los pesos de las conexiones en la red neuronal permanecen sin cambios. Posteriormente, es crucial evaluar si la red neuronal puede resolver nuevos problemas de manera efectiva, incluso aquellos para los cuales ha sido previamente entrenada. Por consiguiente, para validar la red neuronal, se hace necesario utilizar otro conjunto de datos, denominado testeo o conjunto de validación [34].

Cada vez que se realiza una prueba con el conjunto de validación, se incluyen los valores de las variables de entrada y su solución correspondiente. Sin embargo, en este punto, la solución no se revela a la red neuronal. Luego, se compara la solución calculada para cada instancia de validación con la solución conocida. El nuevo ejemplo utilizado para la validación se llama  $E_u$  y su salida correcta correspondiente es  $A_u$  (donde "u" representa la incógnita en inglés). El desafío consiste en determinar cuándo se debe considerar como correcta la salida de la red neuronal [34].

### 1.3.9 Modelos de inteligencia artificial

Los modelos de inteligencia artificial son programas y algoritmos informáticos que utilizan la inteligencia artificial para tomar decisiones de forma autónoma. Estos modelos son representaciones de las capacidades de la mente humana y se crean con el propósito de automatizar acciones que requieren inteligencia. Para lograr esto, es fundamental contar con una base de experiencia que permita abordar problemas desconocidos previamente, lo cual es una característica clave de la inteligencia. La máquina puede adaptar este principio mediante el análisis de datos. Existen diversas aplicaciones y enfoques para dotar a los programas y algoritmos de la capacidad de realizar acciones de manera inteligente [36].

#### *a. Construcción de modelos*

La inteligencia artificial en la construcción de modelos es un proceso esencial en la implementación de esta tecnología. Consiste en una secuencia de etapas planificadas con el objetivo de convertir datos en conocimientos relevantes y habilidades predictivas. Cada fase desempeña un papel crucial en el desarrollo de sistemas inteligentes capaces de aprender y tomar decisiones autónomas. A continuación, en la Tabla 13 se presentan de manera concisa y técnica los pasos fundamentales en la construcción de un modelo con IA:

Tabla 13. Procesos de construcción de modelos [36].

<b>Procesos de construcción de modelos</b>	<b>Característica</b>
<b>Recolección de datos</b>	Se puede adquirir los datos de diversas fuentes, como la extracción de información de un sitio web, la utilización de una API o la obtención de datos desde una base de datos. También es posible emplear dispositivos específicos para la recopilación automática de datos, así como utilizar información de dominio público. Las posibilidades para la recolección de datos son prácticamente ilimitadas. Aunque este paso pueda parecer evidente, es uno de los que conlleva mayores desafíos y consume una cantidad significativa de tiempo.
<b>Procesar datos</b>	Una vez que se dispone de los datos, es necesario verificar que estos posean el formato adecuado para alimentar el algoritmo de aprendizaje. Resulta prácticamente inevitable llevar a cabo diversas tareas de preprocesamiento antes de poder emplear los datos. Aunque este proceso suele ser más simple en comparación con la fase de recolección, sigue siendo esencial para la preparación efectiva de los datos.

<b>Procesos de construcción de modelos</b>	<b>Característica</b>
<b>Exploración de datos</b>	Una vez que los datos se encuentran disponibles y han sido formateados adecuadamente, se procede a realizar un preanálisis para abordar situaciones de valores faltantes y para identificar patrones que puedan facilitar la construcción del modelo. En esta fase, las medidas estadísticas y los gráficos en dos y tres dimensiones resultan herramientas de gran utilidad para comprender el comportamiento de nuestros datos. Además, en este punto, es posible detectar valores atípicos que podrían ser excluidos de la predicción, así como identificar las características que ejercen una influencia significativa en la capacidad predictiva del modelo.
<b>Entrenamiento del algoritmo</b>	En esta fase, se aplica efectivamente las técnicas de Machine Learning. Se alimenta uno o varios algoritmos de aprendizaje con los datos que se han procesado en las etapas previas. La premisa es que estos algoritmos sean capaces de extraer información valiosa de los datos proporcionados, permitiéndonos realizar predicciones.
<b>Evaluación del algoritmo</b>	En esta fase, se somete a evaluación la información o conocimiento adquirido por el algoritmo durante la fase de entrenamiento previa. Se analiza la precisión de las predicciones del algoritmo y, en caso de no estar satisfechos con su rendimiento, se regresa a la etapa anterior y se continúa con el entrenamiento ajustando algunos parámetros hasta alcanzar un rendimiento aceptable.
<b>Implementación del modelo</b>	En esta etapa final, se somete el modelo al entorno del problema real. En este punto, se evalúa el rendimiento, lo que podría requerir una revisión exhaustiva de todos los pasos anteriores. El modelo determina cómo los datos de entrada de cada solicitante pueden ser empleados para mejorar la predicción del resultado del préstamo. Al identificar y utilizar patrones en el conjunto de entrenamiento, el aprendizaje automático produce un modelo (pensado como una caja negra) que genera predicciones sobre el resultado para cada nuevo solicitante basándose en sus datos específicos.

### ***b. Evaluación de modelos***

Una práctica recomendada implica la partición del conjunto de datos etiquetados en dos subconjuntos: uno destinado al entrenamiento y otro al testeo o validación. La determinación del tamaño de cada conjunto suele ser una decisión importante. Aunque no hay una regla universal para establecer la proporción exacta, es común encontrar problemas en los cuales alrededor del 90% se asigna para el conjunto de entrenamiento y aproximadamente el 10% para el conjunto de test [33].



- El conjunto de entrenamiento se utiliza con el propósito de desarrollar un modelo de clasificación o regresión mediante el aprendizaje de patrones presentes en los datos [33].
- El conjunto de test se emplea con la finalidad de realizar predicciones sobre el valor de instancias específicas mediante la aplicación del modelo previamente entrenado [33].

Una vez definido el objetivo para aplicar Machine Learning, el primer paso implica examinar la disponibilidad de datos, su cantidad y características. Idealmente, se busca un conjunto de datos extenso o al menos de tamaño moderado. La selección de características busca reflejar la estructura del problema y equilibrar la diversidad entre las características. Dado que los datos siempre contienen ruido y comportamientos anómalos, es crucial encontrar un conjunto de datos que refleje con precisión la realidad del problema [33].

La organización de los datos implica la división en dos grupos: conjunto de entrenamiento y conjunto de testeo, con una proporción recomendada del 80%-20%. En el tratamiento del conjunto de entrenamiento, se aplica la técnica de validación cruzada, dividirlo en varios conjuntos de entrenamiento y prueba para sucesivas etapas. El diseño del predictor implica seleccionar un algoritmo o función de predicción que se aplicará al conjunto de entrenamiento. En este diseño, se equilibra la precisión de los resultados con otras características deseables, como se muestra en la Figura 17 [33].

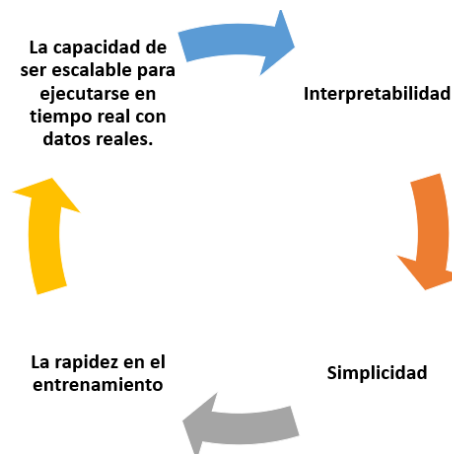


Figura 17. Atributos preferibles del modelo [33].

Finalmente, se aplica la función de predicción a los datos de entrenamiento utilizando técnicas de validación cruzada para refinar el modelo hasta lograr un error aceptable. El resultado final puede ser la combinación de varios predictores previamente probados [33].

Una vez obtenido el modelo, se procede a evaluar su precisión mediante la aplicación de la función de predicción a los datos de prueba, midiendo el error resultante. Si el error excede el umbral considerado aceptable, se inicia un proceso de refinamiento de la función de predicción. Los resultados se clasifican utilizando la matriz de confusión, que incluye categorías como se indica en la Tabla 14 [33]:

Tabla 14. Categorías de la matriz de confusión [33].

Categorías	Descripción
<b>Falsos positivos</b>	Se refiere a resultados que son incorrectamente identificados como correctos.
<b>Falsos negativos</b>	Son resultados que son equivocadamente identificados como incorrectos.
<b>Positivos verdaderos</b>	Son resultados correctamente identificados como positivos
<b>Negativos verdaderos</b>	Son resultados correctamente identificados como negativos.



Figura 18. Matriz de confusión [33].

Basado en la clasificación de la Figura 18, se determina dos indicadores fundamentales [33]:

- **Sensibilidad:** También conocida como “recall” o tasa de verdaderos positivos, se define como el cociente entre los verdaderos positivos y el total de casos positivos reales. [33].
- **Especificidad:** Se calcula como el cociente entre los verdaderos negativos y el total de casos negativos reales. [33].

Estos indicadores, Sensibilidad y Especificidad, son cruciales para evaluar la calidad de un predictor mediante la construcción de la Curva ROC (Característica Operativa del Receptor). Esta curva relaciona la sensibilidad con (1-especificidad). Un predictor altamente efectivo se reflejará en un Área bajo la Curva (AUC) cercana a uno, indicando un rendimiento excepcional. A medida que el AUC se acerca a 0.5, la utilidad del predictor disminuye. Al disponer de múltiples modelos, la comparación de los valores del AUC facilita la elección del más adecuado [33].

En el proceso de diseño del predictor, surge el riesgo conocido como "Sobreajuste" (Overfitting en inglés), relacionado con la capacidad de la función de predicción para clasificar eficazmente los datos de entrenamiento, siendo demasiado específica y susceptible a incorporar ruido. Al probar la función en datos de testeo, es posible encontrarse con que versiones anteriores de la función muestran un mejor rendimiento. El AUC se presenta como un indicador útil para comparar diferentes iteraciones de la función de predicción, la Tabla 15 nos muestra los indicadores de las métricas de rendimiento del modelo [33]:

Tabla 15. Métricas de rendimiento del modelo [33].

Métricas	Descripción
<b>Precisión</b>	Evalúa la proporción de instancias correctamente clasificadas como positivas respecto al total de instancias positivas. Esta métrica resulta valiosa cuando los falsos positivos tienen un impacto significativo.
<b>F1 Score</b>	Representa la media armónica del recall y la precisión, proporcionando un equilibrio ponderado entre ambas métricas. Esta medida se vuelve especialmente útil en situaciones donde existe un desequilibrio notorio en la distribución de clases.
<b>Exactitud o Accuracy</b>	Cuantifica la proporción de instancias clasificadas correctamente (positivas y negativas) respecto al total de instancias. Se recomienda su uso cuando las clases se encuentran balanceadas en el conjunto de datos.

<b>Métricas</b>	<b>Descripción</b>
<b>Coefficiente Kappa de Cohen (k)</b>	Este coeficiente relaciona la concordancia observada con la concordancia esperada por azar, proporcionando una medida de la fiabilidad del modelo.
<b>Desviación Media.</b>	Se calcula como la media de la suma de las desviaciones absolutas, es decir, de la diferencia entre el valor y la media se calcula el valor absoluto para su resultado
<b>Error Cuadrático Medio</b>	Se define como la suma de las diferencias al cuadrado entre los valores reales y los predichos.

## **1.4 Objetivos**

### **1.4.1 Objetivo general**

Implementar un sistema de detección de emociones mediante el análisis de indicadores faciales empleando inteligencia artificial.

### **1.4.2 Objetivos específicos**

- Analizar los indicadores faciales que permiten identificar las emociones.
- Determinar algoritmos de inteligencia artificial para la identificación de rostros y determinación de emociones.
- Diseñar un sistema electrónico detector de emociones que incorpore una interfaz interactiva para los usuarios.

## **CAPÍTULO II. METODOLOGÍA**

### **2.1 Materiales**

En el ámbito de la investigación destinado al desarrollo del presente proyecto, se utilizaron diversos dispositivos equipo para llevar a cabo el proceso de captura de frames en tiempo real y el procesamiento de las imágenes. Se utilizó una cámara web de alta resolución para la obtención de imágenes, complementada con la microcomputadora NVIDIA Jetson Nano. La elección de esta última se basa en las características que proporciona para un procesamiento de datos eficiente y rápido, contribuyendo así a la agilidad del conjunto del sistema.

Para el manejo de datos se empleó el lenguaje de programación Python, dentro del cual se encuentra el modelo entrenado en Google Colab. Este entorno proporciona facilidades de programación modular y ofrece la opción de utilizar su librería TensorBoard. A partir de los resultados obtenidos, se establece una conexión con la base de datos implementada en MySQL y alojada en el servidor de 000WebHosting. En el mismo entorno de alojamiento en la nube, se ha desarrollado la interfaz administrativa destinada a la gestión de resultados.

### **2.2 Métodos**

#### **2.2.1 Modalidad de la investigación**

El presente proyecto es una investigación aplicada, debido a que se buscará dar una solución al problema planteado mediante la aplicación práctica de conocimientos para el desarrollo de un sistema electrónico utilizando técnicas de inteligencia artificial.

La investigación es de tipo bibliográfica, ya que se fundamenta en la recopilación de información proveniente de libros, artículos científicos, datasheets de los dispositivos, publicaciones en internet y tesis relacionadas con la implementación del sistema.

La investigación se considera de campo, porque es esencial consultar con profesionales de la salud y realizar pruebas en una variedad de individuos en diferentes ámbitos

sociales para ajustar el sistema y asegurarse de que sea un dispositivo de alta confiabilidad.

### 2.2.2 Población y muestra

El presente proyecto de investigación incluyó la implementación de un prototipo libre, por lo tanto, se puede llevar a cabo en cualquier ámbito, y no requiere población y muestra.

### 2.2.3 Recolección de información

Para la obtención de información se empleó los distintos repositorios de universidades del país, libros, revistas de difusión científica en línea, y demás bases de datos que sean confiables con la finalidad que puedan aportar para el desarrollo del proyecto, como se explica en la Tabla 16:

Tabla 16. Recolección de información.

<b>Criterio</b>	<b>Descripción</b>
¿Para qué?	Determinar el estado en tiempo real de las emociones
¿De qué personas u objetos?	Para personas en varios ámbitos
¿Cómo?	A través del uso de técnicas de visión artificial
¿Con qué?	Con un sistema de detección de emociones en tiempo real
¿Cuándo?	En el período de octubre 2023 a marzo 2024
¿Dónde?	En ámbitos sociales y de salud

### 2.2.4 Procesamiento y análisis de datos

Para el procesamiento y análisis de datos dentro del marco del proyecto de investigación, se ha definido un conjunto de procedimientos metodológicos. Estos procedimientos abarcan:

- Recolección detallada de información de diversas fuentes.
- Revisión de las alternativas de resolución presentadas con anterioridad en el campo de investigación.

- Análisis de la fundamentación teórica acerca de la conceptualización de las variables de investigación.
- Interpretación de la información más relevante y útil que contribuyan para el desarrollo del proyecto de investigación.

## **CAPÍTULO III. RESULTADOS Y DISCUSIÓN**

### **3.1 Análisis y discusión de los resultados**

Para el sistema de reconocimiento de emociones se utilizó una Jetson Nano como plataforma, se centra en la detección y análisis de expresiones faciales para inferir estados emocionales en tiempo real. Este sistema emplea algoritmos de inteligencia artificial para procesar datos visuales, evaluando características como gestos faciales, movimientos oculares y otras señales no verbales que indican emociones específicas.

La fase inicial del desarrollo involucró la investigación y documentación exhaustiva de los indicadores visuales asociados con diversas emociones humanas. Esta fase fue crucial para identificar y definir los parámetros relevantes que posteriormente fueron utilizados en la implementación de los algoritmos de reconocimiento. La Jetson Nano, con su capacidad de procesamiento eficiente, desempeñó un papel fundamental en garantizar una ejecución rápida y en tiempo real de estos algoritmos.

Durante la fase de entrenamiento, los algoritmos demostraron una eficacia notable, alcanzando una precisión de clasificación del 0.92. Estos resultados validan la capacidad del sistema para identificar con precisión y clasificar estados emocionales dentro del conjunto de datos utilizado en la fase de entrenamiento.

Es crucial destacar que, aunque los resultados son prometedores, la eficacia del sistema puede variar en entornos reales. Por lo tanto, se recomienda la realización de pruebas adicionales en situaciones del mundo real para garantizar la fiabilidad y la adaptabilidad del sistema en diversos contextos. Este enfoque en la evaluación y validación del rendimiento contribuye a la versatilidad y utilidad del sistema de reconocimiento de emociones.

### **3.2 Desarrollo de la propuesta**

En el desarrollo del proyecto, se llevó a cabo una modalidad investigativa, bibliográfica y de campo, enfocada en la definición y documentación de métodos para adquirir indicadores faciales de las emociones. La implementación del sistema de



reconocimiento de estos indicadores se desarrolló a través de un modelo programado en Python, aprovechando su versatilidad y la disponibilidad de librerías con inteligencia artificial. La elección de la librería OpenCV para el procesamiento de imágenes añadió una capa eficiente al sistema, gracias a la variedad de bibliotecas y librerías específicas para el tratamiento de datos visuales o frames .

### 3.2.1 Definición de Emociones para la propuesta

Las emociones desencadenan una señal que facilita la interpretación de situaciones comunes en la vida, una característica que ha perdurado genéticamente en nuestra especie. En esencia, las 6 emociones básicas se consideran una respuesta adaptativa para enfrentar diversos escenarios contextuales es por esta razón que se han determinado para su estudio, incluyendo una emoción neutra para un mejor rendimiento del sistema, a continuación, en la Tabla 17 se presentarán las emociones que se trataron en este sistema [37]:

Tabla 17. Emociones propuestas [37].

<b>Emoción</b>	<b>Concepto</b>
<b>Felicidad</b>	Se destaca por ser una de las más positivas, estando directamente relacionada con el placer y la satisfacción. Esta emoción surge, por ejemplo, cuando se logran objetivos personales o se reduce el malestar.
<b>Tristeza</b>	Es quizás una de las más negativas. Representa una disminución en el estado de ánimo y una notable reducción en la actividad cognitiva y conductual.
<b>Ira</b>	Implica una sensación de poder y, a menudo, surge como respuesta a la discrepancia entre la situación actual y el estado deseado. Cuando las personas sienten una mentalidad de concesión amenazada, esto puede llevar a la frustración.
<b>Miedo</b>	Es una emoción ampliamente estudiada en humanos y animales. Se caracteriza como un estado emocional negativo o aversivo, con un alto poder de activación que promueve la evitación de situaciones potencialmente peligrosas. Experimentar miedo lleva a una sensación de nerviosismo, junto con preocupación por la seguridad y la salud.
<b>Sorpresa</b>	La sorpresa se define como una reacción ante algo inesperado, novedoso o extraño. Implica que, en el momento en que ocurre el estímulo, las personas no habían previsto ni planeado dicha situación. La experiencia subjetiva que sigue a menudo incluye sensaciones de incertidumbre y una sensación de quedarse en blanco.
<b>Disgusto</b>	Se considera una de las emociones básicas. Representa una respuesta ante la posibilidad, real o imaginaria, de evitar la ingestión de sustancias peligrosas que podrían ser contaminantes. En términos subjetivos, se relaciona con un fuerte desagrado y una aversión obvia hacia estímulos ofensivos.
<b>Neutro</b>	Son aquellas que carecen de una cualidad claramente agradable o desagradable, son emociones que no se inclinan ni hacia lo negativo ni hacia lo positivo. Aunque comparten ciertas características tanto con las emociones positivas, en términos de su breve duración temporal, como con las emociones negativas, en cuanto a la considerable movilización de recursos que provocan.

### *a. Detección del rostro*

- **Identificación de elementos clave.** La tarea de localizar los elementos faciales clave presenta desafíos similares a la detección de rostros humanos, pero con algunas particularidades dignas de mencionar. Por un lado, se simplifica la tarea debido a la garantía de que se encontrará una cara con sus elementos esenciales, como los ojos, la nariz y la boca. Sin embargo, surgen dificultades adicionales al trabajar con objetos de menor tamaño debido a la variabilidad en la apariencia de estos componentes [38].

Existen varias dificultades que pueden surgir al intentar ubicar los elementos faciales clave, entre las cuales se destaca en la Tabla 18 [38]:

Tabla 18. Problemas en la detección de elementos faciales [38].

<b>Dificultades</b>	<b>Descripción</b>
<b>Escasa resolución</b>	Cuando el tamaño de las caras se reduce al mínimo detectable, la identificación de los componentes faciales de manera individual se torna prácticamente imposible. Por ejemplo, en una imagen de cara de tan solo 24x30 píxeles, los ojos pueden ocupar no más de 3x3 píxeles. En estas circunstancias, se hace evidente que la localización solo puede llevarse a cabo considerando la estructura global de la cara.
<b>Expresión facial</b>	Las expresiones faciales tienen un impacto considerable en la apariencia del rostro, especialmente cuando se analizan los elementos faciales de forma independiente. Esta situación es frecuente y puede complicar la tarea.
<b>Oclusión y elementos adicionales</b>	Una cara con oclusión parcial puede ocultar parcial o totalmente algunos componentes, lo que significa que la suposición inicial de "existen dos ojos, una nariz y una boca" no siempre se cumple. La oclusión puede deberse a elementos faciales como gafas, bigote o barba, a objetos externos superpuestos o incluso a giros considerables de la cara.
<b>Sombras</b>	Las sombras pueden hacer que los ojos o la boca parezcan simples manchas de píxeles oscuros, incluso en imágenes de alta resolución. En muchos casos, las sombras son causadas por la propia estructura de la cara.
<b>Estructura facial</b>	También es importante considerar la estructura inherente del rostro humano. Los diversos componentes faciales no aparecen en ubicaciones aleatorias; deben seguir una estructura coherente, como la disposición de los ojos sobre la boca, las cejas sobre los ojos y la formación de un triángulo isósceles entre los ojos y la boca.

El reconocimiento de expresiones faciales se lleva a cabo en dos etapas. En la primera fase, se realiza la detección y normalización de la imagen facial, mientras que la segunda etapa se enfoca en el reconocimiento de la expresión facial utilizando la imagen facial previamente normalizada como entrada. La detección de rostro implica la localización de las coordenadas de la cara en una imagen, seguida por un proceso de normalización que asegura una detección precisa [38].

- **Adquisición de indicadores faciales.** Las expresiones faciales de las emociones parecen constituir el elemento central de las interacciones sociales que define a todos los grupos humanos, quizás actuando como un rasgo conductual distintivo que identifica a nuestra especie, es importante tener en cuenta que estas características son generalizaciones y pueden variar según el contexto cultural, individualidad y otros factores. que se presenta a continuación en la Tabla 19 [7].

Tabla 19. Indicadores faciales de las emociones [7].

Características	Felicidad	Tristeza	Ira	Miedo	Sorpresa	Disgusto
Ojos	Abiertos o lagrimas	Hacia abajo y lagrimas	Mirada profunda	abiertos	Muy abiertos	Estrechos
Boca	Sonrisa	Cerrada	Cerrada	Abierta	Poco abierta	Abierta
Pupilas	Brillantes	Dilatadas	Dilatadas			
Cejas	Levantadas	Juntas	Bajas y contraídas al mismo tiempo	Levantadas y contraídas	Levantadas, colocándose curvas y elevadas	Bajas y contraídas al mismo tiempo
Labios	Abiertos	Comisura se inclina hacia abajo, pueden temblar	Apretado abiertos, tensos y en forma cuadrangular	Tensos o ligeramente contraídos	Elevados	Elevados
Frente	Sin arrugas	Arrugas en el centro	Arrugas en el centro y no extendidas	Arrugas en el centro	Arrugas en el centro	Arrugas en el centro
Mejillas	Elevadas	Elevadas		Elevadas		
Parpado inferior	Arrugas debajo	Arrugas debajo	Tenso, puede estar levantado o no	Tenso y alzado	Tenso y alzado	Tenso o levantadas
Parpado superior	Sin arrugas	Sin arrugas	Tenso	levantado	Sin arrugas	Tenso y levantadas

- **Características de las emociones**

Es importante tener en cuenta que cada emoción se manifiesta de acuerdo con las características indicadas en los gráficos de la Tabla 20 [7].

Tabla 20. Características de las emociones definidas [7].

Emoción	Emoción	Característica
Felicidad		<p>Sensación placentera de satisfacción o bienestar, sonriendo tirando hacia arriba las comisuras de los labios, contrayendo los músculos grandes orbitales alrededor de los ojos.</p>
Tristeza		<p>Sentimiento de desdicha o melancolía, los párpados caídos, la reducción de las esquinas de la boca, labios fruncidos, los ojos bajos.</p>
Ira		<p>Frecuentemente expresa el sentimiento de antagonismo hacia una persona o cosa surge tras percibir una afrenta o una ofensa, La reducción de cejas, apretar y estrechar los labios, los ojos mirando fijamente, apretando los párpados inferiores, con menos frecuencia, empujando la mandíbula hacia delante.</p>
Miedo		<p>Sentimiento de inquietud generado por la percepción de peligro, amenaza o la posibilidad de sufrir daño, Levantando las cejas tensando los párpados inferiores, que se extiende horizontalmente labios, la boca ligeramente abierta.</p>
Sorpresa		<p>Sentimiento de incomodidad o sorpresa frente a un acontecimiento inesperado, Levantando las cejas altas (que puede causar arrugas en la frente), abriendo los ojos como platos, dejando caer la mandíbula.</p>
Disgusto		<p>Desagrado o rechazo generado por algo considerado ofensivo o repulsivo. La reducción de cejas, curvando el labio superior, arrugando la nariz</p>

En base a la Tabla 19 y Tabla 20, podemos identificar algunas similitudes y diferencias entre las características asociadas con diferentes emociones:

- Similitudes entre la felicidad y la sorpresa:

Ambas emociones comparten características como los ojos abiertos y la presencia de una sonrisa en la boca, además, en ambas emociones, las pupilas suelen estar dilatadas, lo que puede reflejar una sensación de excitación o alegría.

- Similitudes entre el miedo y la sorpresa:

Ambas emociones comparten la característica de tener los ojos muy abiertos, lo que indica una respuesta de alerta o sorpresa ante una situación inesperada.

- Diferencias entre la tristeza y la ira:

Mientras que la tristeza se caracteriza por los ojos hacia abajo y con lágrimas, la boca cerrada y las cejas juntas, la ira se manifiesta con una mirada dura o prominente, la boca cerrada o tensa y las cejas bajas y contraídas al mismo tiempo, estas diferencias sugieren expresiones faciales distintas que reflejan emociones opuestas: la tristeza está asociada con sentimientos de pesar o desánimo, mientras que la ira implica un estado de enojo o frustración.

- Diferencias entre el asco y las demás emociones:

El asco se distingue por características únicas como los ojos estrechos y la boca abierta, lo que refleja una reacción de repulsión o desagrado ante algo desagradable o repulsivo, esta expresión facial es diferente a las otras emociones consideradas, lo que indica que el asco tiene una manifestación física distintiva.

Mientras que algunas emociones comparten ciertas características faciales, otras presentan expresiones únicas que las distinguen claramente. Es importante considerar el contexto y la combinación de señales emocionales para interpretar con precisión el estado emocional de una persona.

### 3.2.2 Tecnologías utilizadas

#### a. *Microcontrolador*

Una computadora de placa única (SBC) es un sistema computacional integral en el cual una única placa de circuito integra memoria, dispositivos de entrada/salida, un microprocesador y todas las demás funcionalidades esenciales. A diferencia de una computadora personal, la SBC no se apoya en expansiones para funciones adicionales. La utilización de una computadora de placa única resulta en una disminución del costo total del sistema al reducir la cantidad de placas de circuito, conectores y circuitos controladores [39].

- **Raspberry Pi 4.** El Raspberry Pi 4 que se observa en la Figura 19 se configura como una computadora de placa única (SBC) con capacidades adecuadas para aplicaciones de inteligencia artificial (IA). Integra un procesador de 64 bits Broadcom BCM2711 que consta de cuatro núcleos Cortex-A72, operando a una frecuencia de 1.5 GHz. Con opciones de memoria RAM de 2, 4 o 8 GB, este modelo proporciona un rendimiento adaptable para actividades relacionadas con la IA. Su GPU VideoCore VI incorporada facilita el procesamiento gráfico y la aceleración de aplicaciones de IA. Además, incorpora puertos USB 3.0, puertos Ethernet Gigabit, Wi-Fi y Bluetooth, simplificando la conexión de periféricos y las comunicaciones en red requeridas para despliegues de IA. Este dispositivo es comúnmente empleado en proyectos de IA a pequeña escala, como el reconocimiento de objetos, la clasificación de imágenes y tareas de aprendizaje automático de nivel inicial [39].



Figura 19. Raspberry Pi 4 [39]

- ***NVIDIA Jetson Nano.*** El NVIDIA Jetson Nano Development Kit que se observa en la Figura 20 se configura como una computadora de inteligencia artificial diseñada específicamente para creadores, estudiantes y desarrolladores, con el objetivo de brindar la capacidad de la inteligencia artificial contemporánea en una plataforma de fácil manejo y consumo energético eficiente. Este kit proporciona el rendimiento y las capacidades necesarias para llevar a cabo tareas avanzadas relacionadas con la inteligencia artificial. Se trata de una computadora compacta y potente que permite la ejecución simultánea de diversas redes neuronales, abordando aplicaciones como clasificación de imágenes, detección de objetos, segmentación y procesamiento de voz. Gracias a su elevado poder de procesamiento y su enfoque en la inteligencia artificial local, el Jetson Nano resulta idóneo para aplicaciones tales como sistemas autónomos, robótica, visión por computadora y análisis de datos en tiempo real [40].



Figura 20. Jetson Nano [40]

- ***Coral Dev Board***

El Coral Dev Board en la Figura 21 representa una computadora de placa única diseñada para facilitar la creación rápida de prototipos e implementación de sistemas integrados con inferencias de aprendizaje automático (ML) en el dispositivo. Incluye el coprocesador Edge TPU, un ASIC diseñado por Google que acelera eficientemente los modelos TensorFlow Lite, logrando realizar 4 billones de operaciones por segundo (TOPS) con un consumo energético de 0.5 vatios por TOPS. Este dispositivo ofrece un sistema integral que combina una computadora de placa única con SoC, capacidades de ML y conectividad inalámbrica, todo ejecutándose en una placa con un derivado de Debian Linux llamado Mendel, lo que permite la ejecución de

herramientas Linux convencionales. Además, es compatible con TensorFlow Lite, eliminando la necesidad de construir modelos desde cero, ya que los modelos de TensorFlow Lite se pueden compilar para ejecutarse en Edge TPU. Cabe destacar que los términos y características de los productos internacionales pueden variar respecto a los locales, incluyendo ajustes, calificaciones de edad e idioma, etiquetado o instrucciones [41].



Figura 21. Coral Dev Board [41].

- **Comparación de microcomputadoras.** A continuación, en la Tabla 21, se muestra la comparación de las microcomputadoras

Tabla 21. Comparación de placas SBC [39], [40], [41] .

Características	Raspberry Pi 4	Jetson Nano	Coral Dev Board
<b>Fabricante</b>	Raspberry	Nvidia	Google
<b>Procesador</b>	Broadcom BCM2711, Cortex-A72 (4 núcleos)	Quad-core ARM Cortex-A57 (4 núcleos)	Quad-core ARM Cortex-A53 (4 núcleos)
<b>Frecuencia del procesador</b>	Hasta 1.5 GHz	Hasta 1.43 GHz	Hasta 1.5 GHz
<b>GPU</b>	VideoCore VI 500 MHz	Maxwell de 128 núcleos	PowerVR GE8300 GPU
<b>Memoria Ram</b>	4GB LPDDR4	4GB LPDDR4	2GB LPDDR4
<b>Procesador de IA</b>	No específico	NVIDIA Maxwell con 128 núcleos CUDA	Edge TPU (TensorFlow Lite)
<b>Puertos USB</b>	2 x USB 3.0, 2 x USB 2.0	4 x USB 3.0, USB 2.0	1 x USB 3.0, 1 x USB 2.0
<b>Salida de video</b>	2 x micro HDMI	HDMI	HDMI
<b>Puertos Ethernet</b>	Gigabit Ethernet	1 x Gigabit Ethernet	1 x Gigabit Ethernet
<b>Almacenamiento</b>	Ranura para tarjeta microSD	Ranura para tarjeta microSD	8GB eMMC, Ranura para tarjeta microSD
<b>Precio</b>	260.00	180.00	150.00



- **Selección del microcomputador.** Tras evaluar detenidamente los requisitos particulares del proyecto, se decidió seleccionar la Jetson Nano como la alternativa más idónea debido a su versatilidad al capturar imágenes resulta especialmente ventajosa para nuestro proyecto de investigación sobre el reconocimiento de emociones, ya que se ajusta de manera eficiente a las exigencias de la visión artificial en este contexto.

#### **b. Cámaras**

La cámara es un instrumento que registra imágenes transformándolas en señales eléctricas, generalmente en forma de señal de vídeo.

- **Cámara web Gazal Sx1023.** La cámara ha sido diseñada para proporcionar una calidad visual y de audio excepcional. Es perfecta tanto para el uso en clases en línea como para reuniones de trabajo a través de plataformas como Zoom, Skype, Meets o Teams, redefiniendo así tus interacciones virtuales. Con una impresionante resolución de video HD 1080p y una lente de 3.6 mm que ofrece un ángulo de visión de 90°, cada detalle cobra vida en tu pantalla. Las cámaras web Gazal que se visualiza en la Figura 22 desempeñan un papel crucial en al ámbito de la inteligencia artificial ya que es una fuente de video de entrada en aplicaciones de visión artificial como el reconocimiento facial y la detección de objetos [42].



Figura 22. Cámara web Gazal Sx1023 [42].

- **Cámara IMX219-77.** La cámara IMX219-77 o también llamada HBVCAM-HPLCC de la Figura 23 está diseñada específicamente para su uso con las placas Jetson Nano. Incorpora un sensor de imagen Sony IMX219 de 8mp, pero se distingue de las cámaras convencionales por su capacidad única para la captura de imágenes en condiciones de baja luminosidad y en entornos sin luz visible. Es notable su utilización en el ámbito de la inteligencia artificial, donde la cámara se emplea de manera extensiva en proyectos de visión por computadora que requieren la captura de imágenes en entornos de baja luminosidad [43].



Figura 23. Cámara HBVCAM-HPLCC

- **Cámara web Logitech C270.** Logitech, una renombrada corporación suiza con alcance global, se caracteriza por su habilidad en el diseño de productos y experiencias que mejoran las rutinas cotidianas de sus usuarios. La compañía introduce una cámara que se adapta de manera integral a la vida diaria del usuario. La simplificación operativa redefine la experiencia de las videoconferencias, gracias a la implementación de tecnología plug and play que obvia la necesidad de configuraciones elaboradas y ajustes detallados. En el contexto de la visión por computadora, las cámaras Logitech es líder como herramientas fundamentales para diversas aplicaciones, mostrando su idoneidad en funciones tales como reconocimiento facial, seguimiento de objetos y análisis de emociones [44].



Figura 24. Cámara web Logitech C270 HD 30FPS [44].

- **Comparación de cámaras.** A continuación, en la Tabla 22, se detallan las características más importantes de las cámaras investigadas:

Tabla 22. Comparación de las cámaras [42] [43] [44].

Características	Cámara web Gazal Sx1023	Cámara IMX219-77	Cámara web Logitech C270
<b>Tamaño</b>	18 x 7,8 x 12,3 cm	24 x 23 x 10,5 mm	22 x 7,7 x 15,2 cm
<b>Resolución</b>	8mp, HD	8mp	3mp
<b>Sensor de imagen</b>	Si	Sony IMX219	No
<b>Ángulo de visión</b>	180 grados	160 grados	180 grados
<b>Visión nocturna</b>	Tecnología RigtLight para condiciones de poca luz	Infrarrojo con luz de relleno	No
<b>Conexión</b>	USB 2.0	USSB 2.0	USB 2.0
<b>Costo</b>	35.00	25.00	15.00

- **Selección de la cámara.** Luego de un análisis detallado de los requisitos específicos del proyecto, se ha decidido optar por la Cámara web Gazal Sx1023 . Esta selección se basa en su habilidad para realizar la obtención de frames y procesamiento de imágenes a velocidades elevadas, su total compatibilidad con el microordenador y satisface los requisitos del proyecto, así como también su resolución HD y su infrarrojo para mejorar el enfoque de la persona.

### c. *Software de programación*

La importancia de considerar la correcta programación, instalación y configuración del software en el dispositivo designado para ejecutar las funciones esenciales está estrechamente relacionada con la plena operatividad del sistema de inteligencia artificial. La IA ha surgido como una disciplina fascinante y prometedora en el ámbito informático. Es fundamental comprender cuándo y cómo utilizar estos lenguajes de manera efectiva para lograr el desarrollo de la inteligencia artificial, en la Tabla 23 se detallan algunos de los lenguajes de programación más comúnmente utilizados en el desarrollo de soluciones de IA, cada uno con sus propias ventajas y desafíos particulares. [45].

Tabla 23. Software de programación [45].

<b>Software de programación</b>	<b>Definición</b>
<b>Python</b>	Python ha consolidado su posición como el principal lenguaje para el desarrollo de aplicaciones de inteligencia artificial gracias a su sintaxis simple y legible, así como a la abundancia de bibliotecas disponibles. Un destacado framework en Python es TensorFlow, desarrollado por Google, que destaca especialmente en la creación y entrenamiento de redes neuronales, cruciales en el ámbito del aprendizaje profundo a adquisición de habilidades en la programación de inteligencia artificial con Python involucra la comprensión de algoritmos como la regresión lineal, clustering y clasificación, facilitando la implementación de aplicaciones de IA de manera eficiente.
<b>Matlab</b>	Es un software de programación construido en C y C++, que incluye Java para ciertas funciones, su arquitectura modular ayuda a simplificar la integración con otros programas en diversos lenguajes. Su enfoque se destaca en aplicaciones científicas como inteligencia artificial, aprendizaje automático y aprendizaje profundo. A pesar de que tanto MATLAB como Python son valiosos en estas áreas, MATLAB destaca por su fortaleza en algoritmos y sistemas complejos. Aunque la curva de aprendizaje es significativa, especialmente para principiantes, el invertir tiempo resulta beneficioso, sobre todo para ingenieros e investigadores, siendo más fácil para aquellos que se encuentran familiarizados con Python, JavaScript, C y C++, o con el software de ingeniería y la ciencia de datos.
<b>R</b>	Dentro del ámbito de la programación de inteligencia artificial, R destaca como un lenguaje de programación de inmediata relevancia, gozando de amplio uso en el campo de la estadística y el análisis. Su prominencia se intensifica en el ámbito de la inteligencia artificial, especialmente para aquellos interesados en desarrollar enfoques enfocados en el análisis de datos y el modelado estadístico. La utilidad de R se manifiesta al programar algoritmos de inteligencia artificial centrados en la minería de datos, regresión logística, análisis de series temporales y visualización de datos, consolidándose como una solución eficiente para la programación de inteligencia artificial con un enfoque estadístico.

Software de programación	Definición
<b>Julia</b>	Es un lenguaje de programación creado especialmente para el análisis numérico y científico, equiparable en rendimiento a C++ y Python. Julia dispone de bibliotecas destacadas como Flux y MLJ que facilitan el desarrollo de modelos de Machine Learning. Una parte esencial es LLVM (Low-Level Virtual Machine) que desarrolló un esquema para que los scripts programados en Julia se compilen en varias plataformas de manera eficiente.
<b>Prolog</b>	El lenguaje de programación lógica o Prolog se especializa en el campo de la inteligencia artificial el cual está enfocado en la inferencia lógica y la toma de decisiones. Este lenguaje utiliza hechos y reglas lógicas para realizar un razonamiento deductivo esencial en sistemas expertos y recomendación. Es ideal para programar inteligencia artificial con énfasis en el razonamiento y toma de decisiones, ya que sus algoritmos se centran en la resolución de consultas lógicas y la manipulación de bases de conocimiento. Aunque Prolog es menos difundido que Python o C++, es una opción valiosa para abordar problemas de inteligencia artificial enfocados en la lógica y el razonamiento.

- **Comparación de software de programación.** A continuación, en la Tabla 24 se realiza una comparación entre los diferentes softwares de programación.

Tabla 24. Comparación de software de programación [45].

PARÁMETROS TECNICOS	MATLAB	R	JULIA	PROLOG	PYTHON
<b>Visión mediante computador</b>	Toolbox de análisis de imágenes	Pocas funciones.	Funciones llamadas: • JuliaCV • ImageFeatu	Pocas funciones.	Bibliotecas llamadas: • OpenCV • scikit-image
<b>Procesamiento de Texto</b>	Toolbox de procesamiento de texto.	Funciones llamadas: • tm • quanteda	Función llamada: • TextAnalys is.jl	Solo funciones lógicas	Bibliotecas llamadas: • NLTK • SpaCy.
<b>Procesamiento de Datos</b>	Toolbox de análisis de datos	Solo funciones de visualización de datos	Paquetes llamados: • DataFrame s.jl • Query.jl	Solo funciones lógicas	Bibliotecas llamadas: • Pandas • NumPy.
<b>Procesamiento de Imágenes</b>	Toolbox de manipulación de Imágenes	Funciones llamadas: • Imager • EBImage	Funciones de Images.jl y JuliaImages	No común	Bibliotecas llamadas: • OpenCV • scikit-image
<b>Aprendizaje Automático</b>	Toolboxes de machine Learning	Funciones llamadas: • Caret MLR	Paquetes llamados: • Flux • ML • Scikitlearn	Biblioteca llamada: PrologCV	Bibliotecas llamadas: • TensorFlow • PyTorch • Scikit-Lear

PARÁMETROS TECNICOS	MATLAB	R	JULIA	PROLOG	PYTHON
<b>Redes Neuronales</b>	Toolbox de Redes Neuronales	Funciones llamadas: • Caret • Neuralnet	Biblioteca llamada: • Flux.jl	Biblioteca llamada: Deeplearning Log	Bibliotecas llamadas: • TensorFlow • PyTorch • Keras
<b>Rendimiento</b>	Eficiente en tareas matemáticas y geométricas	Rendimiento lento, sin embargo, con optimización puede mejorar	Rendimiento cercano a: • C++ • Python	Rendimiento lento, sin embargo, con optimización puede mejorar	Rendimiento competitivo con opción de extender con lenguaje C
<b>Comunidad</b>	Gran comunidad con soporte sólido.	Comunidad en crecimiento y soporte activo en estadística y análisis de datos	Comunidad en crecimiento con soporte activo	Comunidad pequeña y soporte activo en aplicaciones lógicas	Enorme comunidad de desarrolladores de código abierto y amplio soporte

- **Selección de software.** Cuando se aborda la programación de inteligencia artificial, la selección apropiada del lenguaje y algoritmo es esencial para el éxito del proyecto. Python se destaca por su versatilidad y comunidad activa, Matlab destaca por su énfasis en el análisis numérico y su sólida integración con herramientas de visualización, R se centra en analítica y estadística, Julia se caracteriza por su velocidad y escalabilidad, además de contar con una comunidad en crecimiento y librerías especializadas en IA, mientras que Prolog se basa en reglas para la toma de decisiones y modelos lógicos [45].

A partir de las especificaciones de la Tabla 24 de cada software, Python se presenta como la opción más adecuada para satisfacer los requisitos del sistema, principalmente gracias a su capacidad para ejecutar la visión artificial de manera más eficiente en comparación con otros lenguajes de programación. Esto se debe a su enfoque en la inteligencia artificial utilizando de scripts para la ejecución, lo que resulta en tiempos de procesamiento considerablemente más cortos. La característica distintiva de Python se convierte en una ventaja significativa para el sistema de detección, permitiendo el procesamiento de imágenes en un periodo de tiempo considerablemente inferior, además Python cuenta con librerías y bibliotecas ampliamente disponibles en la mayoría de los programas de entrenamiento, facilitando así la implementación y ejecución de modelos con inteligencia artificial [45].

#### ***d. Algoritmos de la red neuronal***

Para el diseño de la red neuronal destinada al sistema de detección de emociones, se han aplicado tres algoritmos de inteligencia artificial.

- ***Local Binary Patterns (LBP)***. Se trata de una metodología descriptiva sencilla pero altamente eficaz en el ámbito de la clasificación de objetos dentro del campo de la visión por computadora. Esta técnica realiza un filtrado de los píxeles circundantes mediante consideraciones específicas, generando un valor binario que sirve como representación distintiva. Se basa en la codificación de la relación entre el píxel central y la intensidad de color de los píxeles adyacentes [46].
- ***FaceNet***. FaceNet ha ganado reconocimiento como un sistema de identificación facial creado por Google, que produce embeddings (incrustaciones) de rostros. Estas incrustaciones se presentan como representaciones numéricas que capturan las características únicas de un rostro, destaca particularmente por su habilidad para generar incrustaciones que permanecen invariantes frente a variaciones en la iluminación, posición y expresión facial [47].
- ***Redes Neuronales convolucionales (CNN)***. Se refieren a arquitecturas de aprendizaje profundo que encuentran aplicación frecuente en el procesamiento de imágenes, visión por computadora y reconocimiento facial. Estas redes están conformadas por capas convolucionales diseñadas para aprender de manera jerárquica las características y patrones presentes en las imágenes. Su eficacia se destaca especialmente en la clasificación de imágenes, y han alcanzado notables avances en el ámbito del reconocimiento facial gracias a su capacidad para adquirir representaciones complejas y su resistencia frente a variaciones [47].

Para utilizar la inteligencia artificial se contó con algoritmos eficientes y precisos que puedan interpretar de manera efectiva las expresiones faciales y otros indicadores no verbales. En la Tabla 25 analizaremos detalladamente estas características en relación con su aplicación en sistemas de detección de emociones, considerando tanto la efectividad de la detección como la viabilidad práctica de implementación.

Tabla 25. Especificaciones de los algoritmos para la red neuronal [46] [47].

Algoritmos	Ventajas	Desventajas
<b>Local Binary Patterns (LBP)</b>	<ul style="list-style-type: none"> <li>• <b>Simplicidad:</b> El algoritmo LBP es relativamente simple de entender e implementar.</li> <li>• <b>Robustez:</b> Funciona bien en condiciones de iluminación variable y cambios de escala.</li> <li>• <b>Eficiencia computacional:</b> Es computacionalmente eficiente y rápido en su ejecución.</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Sensible a la textura:</b> LBP puede no ser tan efectivo en la detección de características en imágenes con texturas complejas.</li> <li>• <b>Sensibilidad al ruido:</b> Puede producir resultados subóptimos en presencia de ruido en la imagen.</li> <li>• <b>Dependiente del tamaño de la ventana:</b> La eficacia de LBP puede depender del tamaño de la ventana seleccionada para el análisis de la imagen.</li> </ul>
<b>FaceNet</b>	<ul style="list-style-type: none"> <li>• <b>Precisión:</b> FaceNet ha demostrado una alta precisión en la identificación y verificación de rostros de varios usuarios simultáneamente.</li> <li>• <b>Invariancia:</b> Es invariante a ciertas transformaciones como rotación, escala y cambio de iluminación.</li> <li>• <b>Capacidad para grandes conjuntos de datos:</b> Es capaz de manejar grandes conjuntos de datos y escalar eficientemente.</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Requiere grandes cantidades de datos:</b> FaceNet puede requerir grandes cantidades de datos para entrenar modelos precisos.</li> <li>• <b>Recursos computacionales:</b> La implementación y ejecución de FaceNet puede requerir recursos computacionales significativos.</li> <li>• <b>Sensible a la calidad de los datos:</b> La calidad de los datos utilizados para entrenar el modelo puede afectar significativamente su rendimiento.</li> </ul>
<b>Redes Neuronales convolucionales (CNN)</b>	<ul style="list-style-type: none"> <li>• <b>Flexibilidad:</b> Las CNN pueden adaptarse a una amplia variedad de tareas de visión por computadora, incluido el reconocimiento facial.</li> <li>• <b>Capacidad de aprendizaje:</b> Son capaces de aprender características relevantes de los datos automáticamente.</li> <li>• <b>Estado del arte:</b> Las CNN han demostrado ser líderes en muchas tareas de visión por computadora, incluido el reconocimiento facial.</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Requiere grandes conjuntos de datos:</b> Las CNN pueden requerir grandes cantidades de datos para entrenar modelos precisos.</li> <li>• <b>Sobrecarga computacional:</b> El entrenamiento y la ejecución de CNN pueden ser computacionalmente costosos, especialmente para arquitecturas profundas.</li> <li>• <b>Explicación limitada:</b> A menudo, las CNN se consideran cajas negras, lo que significa que es difícil interpretar cómo y por qué toman decisiones específicas.</li> </ul>

- **Selección de algoritmos.** Para el desarrollo de la red neuronal de sistema de detección de emociones se integró los algoritmos expuestos en la Tabla 25, la metodología inicial se centrará en la adquisición de imágenes a partir de videos en línea, con el uso de la biblioteca OpenCV para capturar varios frames que servirán como entrada para el sistema donde se aplicará la técnica de FaceNet para el reconocimiento de rostros en los frames o imágenes capturadas, después se utilizará LBP durante las etapas de convolución. La selección de LBP se fundamenta por su



capacidad para describir patrones de texturas, lo cual es primordial para extracción de características o patrones específicos relacionadas con las expresiones faciales de las emociones, y así seguirán capas siguientes de convoluciones más pequeñas que permitirán capturar patrones más específicos y complejos. Por último, la fase de toma de decisiones se llevará a cabo después de las capas CNN con neuronas dense o totalmente conectadas, para entender características profundas y clasificarlos de manera precisa. La combinación de LBP, FaceNet y CNN busca aumentar el rendimiento del sistema de detección de emociones para que su identificación y resultados sean precisos y eficientes [48].

#### *e. Plataformas Web Hosting*

El servicio de hosting o alojamiento web constituye una prestación digital cuya finalidad radica en resguardar la información y contenido asociado a un sitio web en centros de procesamiento de datos eficientes y seguros. Su función es crucial para asegurar la disponibilidad en Internet de un sitio web, al proporcionar el espacio y los recursos esenciales para su almacenamiento [48].

- ***Infinity Free.*** Este servicio de alojamiento gratuito destaca por su rendimiento superior y una velocidad notablemente competitiva. Además, fiel a su denominación, proporciona capacidad para alojar cualquier cantidad de sitios web que requieras [49].
- ***000webhost.*** Free Web Hosting se presenta como otra alternativa gratuita que brinda servicios de alojamiento con un desempeño y capacidad destacados. Aunque esta plataforma dispone de algunos planes de pago su servicio gratuito proporciona las herramientas esenciales para llevar a cabo tareas simples o complejas pero fundamentales además que permite gran integración de recursos y datos [49].
- ***Byethost.*** Byethost se presenta como un servicio de alojamiento web que alberga más de 1 millón de sitios que han sido registrados. Los usuarios pueden compartir sus servidores gratuitos generando una notable capacidad de procesamiento. Además, la combinación de su capacidad de almacenamiento y su panel de control lo convierte en una opción atractiva [49].

- **Comparación de Web Hosting.** Las características de los Web hosting se detallan a continuación en la Tabla 26:

Tabla 26. Comparación de Web Hosting [49].

Web Hosting	Ventajas	Desventajas
<b>Infinity Free</b>	<ul style="list-style-type: none"> <li>• Amplio ancho de banda y almacenamiento de 5GB.</li> <li>• Nombre de dominio o subdominio mediante la plataforma</li> <li>• Sin anuncios en el alojamiento web.</li> </ul>	<ul style="list-style-type: none"> <li>• Las cuentas de correo son limitadas</li> <li>• Base de datos MySQL limitadas</li> <li>• No tiene soporte</li> <li>• Potencia del servidor limitada</li> </ul>
<b>000webhost</b>	<ul style="list-style-type: none"> <li>• Es de uso sencillo</li> <li>• Potencia de servidor alta</li> <li>• Sin anuncios en el alojamiento web</li> <li>• Base de datos MySQL ilimitadas.</li> <li>• Plan de pago para más funciones y recursos.</li> </ul>	<ul style="list-style-type: none"> <li>• Incluye un alojamiento web con una cuenta.</li> <li>• Ancho de banda limitado</li> <li>• Pocas plantillas de diseño</li> </ul>
<b>Byethost</b>	<ul style="list-style-type: none"> <li>• Sin anuncios en el alojamiento web</li> <li>• Acceso a la comunidad</li> <li>• Nombre del subdominio gratuito</li> <li>• Potencia de servidor alta</li> </ul>	<ul style="list-style-type: none"> <li>• Límite de espacio para el contenido de la página.</li> <li>• No disponible para Windows</li> <li>• Las cuentas de correo son limitadas</li> <li>• Base de datos MySQL limitadas</li> </ul>

- **Selección de plataformas Web Hosting.** A partir de las especificaciones de cada web hosting, la selección se inclina hacia 000webhost por diversas razones. En primer lugar, su interfaz de usuario intuitiva y el rendimiento sólido de su servidor ofrecen un entorno de uso fácil y eficiente. La ausencia de anuncios en el alojamiento web y la capacidad de contar con bases de datos MySQL ilimitadas son ventajas destacadas. Aunque presenta ciertos inconvenientes, como un ancho de banda limitado y un número reducido de plantillas de diseño proporciona flexibilidad y escalabilidad adaptándose a las necesidades específicas para el desarrollo del proyecto.

#### **f. Herramientas de desarrollo de entrenamiento**

Las herramientas de inteligencia artificial ofrecen soporte a los diseñadores y desarrolladores para automatizar diversas tareas, mejorar la generación de contenido, optimizar los procesos de prueba y agilizar de manera integral los pasos de desarrollo

e implementación. Su objetivo es facilitar a los desarrolladores y diseñadores web la maximización de las oportunidades que brinda la tecnología [50].

- **TensorFlow Playground.** TensorFlow Playground, desarrollada por Google, proporciona un entorno interactivo en línea para la exploración y comprensión de las características fundamentales de las redes neuronales y el aprendizaje profundo. Esta herramienta permite a los usuarios interactuar con los parámetros de modelos, arquitecturas de red y observar de manera visual cómo dichos ajustes influyen en la versatilidad del modelo. Aunque su enfoque principal es educativo, TensorFlow Playground ofrece una interfaz intuitiva que ayuda a asimilar los principios esenciales de TensorFlow y del aprendizaje automático [51].

- **JupyterLab.** JupyterLab representa una interfaz de usuario basada en la web diseñada para la elaboración y desarrollo de documentos interactivos en el formato de Jupyter Notebooks. Ampliando las funcionalidades existentes en los Jupyter Notebooks, esta interfaz aporta una estructura más flexible y robusta. La capacidad de Jupyter Lab para organizar actividades en entornos más organizados se refleja en su capacidad para abrir simultáneamente varios notebooks y editores de código en una única interfaz. Al admitir una amplia gama de lenguajes de programación y ofrecer la posibilidad de integrar extensiones brinda una experiencia de desarrollo enriquecedora, especialmente aplicable a proyectos de machine learning y análisis de datos [52].

- **Amazon SageMaker.** Amazon SageMaker se presenta como una solución integral para el aprendizaje automático, gestionando de manera completa este proceso. Este servicio facilita a los desarrolladores y analistas de datos la creación y refinamiento eficientes de modelos de aprendizaje automático, permitiendo su despliegue directo en un entorno alojado y preparado para su uso inmediato. Incluye una instancia integrada de bloc de notas de Jupyter para acceder sin complicaciones a los conjuntos de datos y realizar investigaciones y análisis sin la necesidad de gestionar servidores. SageMaker no solo ofrece algoritmos de aprendizaje automático convencionales, sino que los optimiza para ejecutarse eficientemente en entornos distribuidos, especialmente adecuados para conjuntos de datos de gran escala. Además, con compatibilidad para algoritmos propios y frameworks nativos, SageMaker brinda

opciones de entrenamiento distribuido y flexibles que se adaptan a los flujos de trabajo específicos. La implementación de modelos se simplifica, permitiendo un lanzamiento seguro y escalable desde SageMaker Studio o la consola de SageMaker [53].

- **Colaboratory (Colab).** Colaboratory, comúnmente denominado Colab, representa un entorno de desarrollo colaborativo que se basa en Jupyter Notebooks y opera en la infraestructura en la nube proporcionada por Google. Al proveer acceso gratuito a recursos computacionales, Colab habilita a los usuarios para la creación y ejecución interactiva de código en Python. La integración sin complicaciones con Google Drive facilita tanto el almacenamiento como el intercambio de códigos o notebooks, mientras que su capacidad de compilación por partes facilita una ejecución minuciosa del código. La destacada accesibilidad, la participación de la comunidad y el respaldo de Google consolidan a Colab como una elección predilecta para el desarrollo y entrenamiento de modelos de inteligencia artificial [54].

- **Comparación de herramientas de entrenamiento.** Se realizó un análisis exhaustivo de las características fundamentales de las herramientas destinadas al desarrollo del entrenamiento de redes neuronales. En la Tabla 27 se presenta de manera detallada las propiedades más significativas:

Tabla 27. Comparación de herramientas de entrenamiento [51] [52] [53] [54].

<b>Características</b>	<b>TensorFlow Playground</b>	<b>JupyterLab</b>	<b>Amazon SageMaker</b>	<b>Colaboratory (Colab)</b>
<b>Accesibilidad</b>	Online	En la nube	En la nube	En la nube de Google drive
<b>Entorno de programación</b>	Interactivo	Interactivo y enlaces para Jupyter Notebooks en la nube	Interactivo e integra servicios de AWS	Interactivo y enlaces para Jupyter Notebooks en la nube
<b>Librerías</b>	TensorFlow	Varias	Varias	Varias
<b>Colaboración</b>	No	Si, en tiempo real	Si, en anexos y documentación	Si, en tiempo real
<b>Interfaz gráfica</b>	No	No	No	TensorBoard (Ayuda a visualizar tu modelo y rendimiento)

Características	TensorFlow Playground	JupyterLab	Amazon SageMaker	Colaboratory (Colab)
Soporte para Frameworks	No	Si	Si	Si
Integración de Hardware	No	No	Si	No
Usabilidad	Fácil	Fácil	Difícil	Fácil
Comunidad	Grande y soporte activo	Pequeña y soporte poco activo	Grande y soporte activo	Enorme y el soporte es extenso debido a desarrolladores colaboradores entre la comunidad
Costo	Gratuito	Gratuito	El pago depende del uso de servicios y recursos	Gratuito, opciones de pago para recursos adicionales

- **Selección de las herramientas de desarrollo de entrenamiento.** A partir de las especificaciones de cada herramienta de desarrollo de entrenamiento, Colaboratory (Colab) sobresale como la opción preferida por varias razones técnicas: que se fundamenta en su accesibilidad, facilidad de uso y en una comunidad activa que ofrece un valioso respaldo y estimula el intercambio de conocimientos, además se distingue por su capacidad única de compilación por partes, simplificando la identificación y corrección de errores en el proceso de desarrollo y entrenamiento. La plataforma también incorpora TensorBoard, una herramienta potente para visualizar detalladamente gráficos de error y rendimiento del modelo. Estas características, combinadas consolidan a Colab como una elección robusta para el desarrollo y entrenamiento eficaz del modelo de inteligencia artificial para la detección de emociones.

### 3.2.3 Arquitectura del sistema

A continuación, en la Figura 25 se detalla la arquitectura del sistema:



Figura 25. Arquitectura del sistema.

**a. Etapa 1: Adquisición y procesamiento de datos**

- Adquisición.** El sistema se activará cuando un usuario intente acceder a la interfaz. Para ello, deberá realizar un proceso de registro y posteriormente iniciar sesión para acceder al sistema. Una vez dentro, se le proporcionará la capacidad de utilizar cualquier cámara disponible la determinación de emociones. Es importante destacar que esta funcionalidad puede llevarse a cabo tanto en tiempo real como en línea, permitiendo una experiencia de usuario versátil y adaptable a diferentes contextos de aplicación.
- Procesamiento.** Una vez que la cámara captura las imágenes del video en tiempo real, estos datos visuales son transmitidos al microcomputador Jetson Nano, el propósito de esta etapa es realizar una manipulación detallada de la información visual, extrayendo los indicadores faciales característicos presentes en la imagen. Este proceso implica la aplicación de técnicas avanzadas de procesamiento digital de imágenes con el fin de obtener de manera precisa y eficiente los rasgos distintivos asociados a la expresión facial. La capacidad de Jetson Nano para llevar a cabo este análisis en tiempo real garantiza una respuesta dinámica y ágil en el reconocimiento de emociones, contribuyendo así a la robustez y eficacia del sistema en su conjunto.

### ***b. Etapa 2: Entrenamiento en la red Neuronal***

- ***Entrenamiento.*** En la fase de entrenamiento, una vez que la imagen ha sido preparada, se somete a un proceso avanzado que implica el empleo de técnicas sofisticadas de inteligencia artificial, destacando el uso de redes neuronales. Este modelo de aprendizaje profundo exhibe la habilidad de identificar de manera precisa los indicadores faciales específicos asociados a cada emoción predefinida. Durante este proceso, la red neuronal se ajusta y perfecciona mediante la exposición a un conjunto diverso y representativo de datos de entrenamiento, permitiendo así que el sistema adquiera la capacidad de generalizar y reconocer patrones complejos de expresiones faciales con alta precisión para contribuir significativamente a la efectividad y confiabilidad del sistema en la tarea de reconocimiento emocional.

### ***c. Etapa 3: Almacenamiento y visualización***

- ***Almacenamiento.*** Con el propósito de facilitar su procesamiento y aplicación práctica, el sistema se implementa en un entorno de web hosting. Al desplegar el sistema en la nube, se aprovechan las ventajas como la escalabilidad y accesibilidad remota, permitiendo así un tratamiento ágil y centralizado de la información, también contribuye a la disponibilidad del sistema, garantizando su accesibilidad desde diversos dispositivos y ubicaciones geográficas.

- ***Visualización.*** En la etapa de visualización, se presenta el estado emocional de la persona en tiempo real. Este estado emocional se visualiza a través de una interfaz específica que indica si la persona continúa en línea junto con su historial de emociones y gráficas, lo que permite la evaluación de sus emociones. Esta información recopilada estará disponible para un usuario administrativo autorizado lo que posibilita la verificación de los usuarios registrados y las emociones capturadas. La implementación de esta estructura de visualización y almacenamiento centralizado contribuye al monitoreo continuo y la evaluación de datos acumulados.

### ***a. Diagrama de bloques***

En la Figura 26 se visualiza se detallan los bloques del sistema implementado:

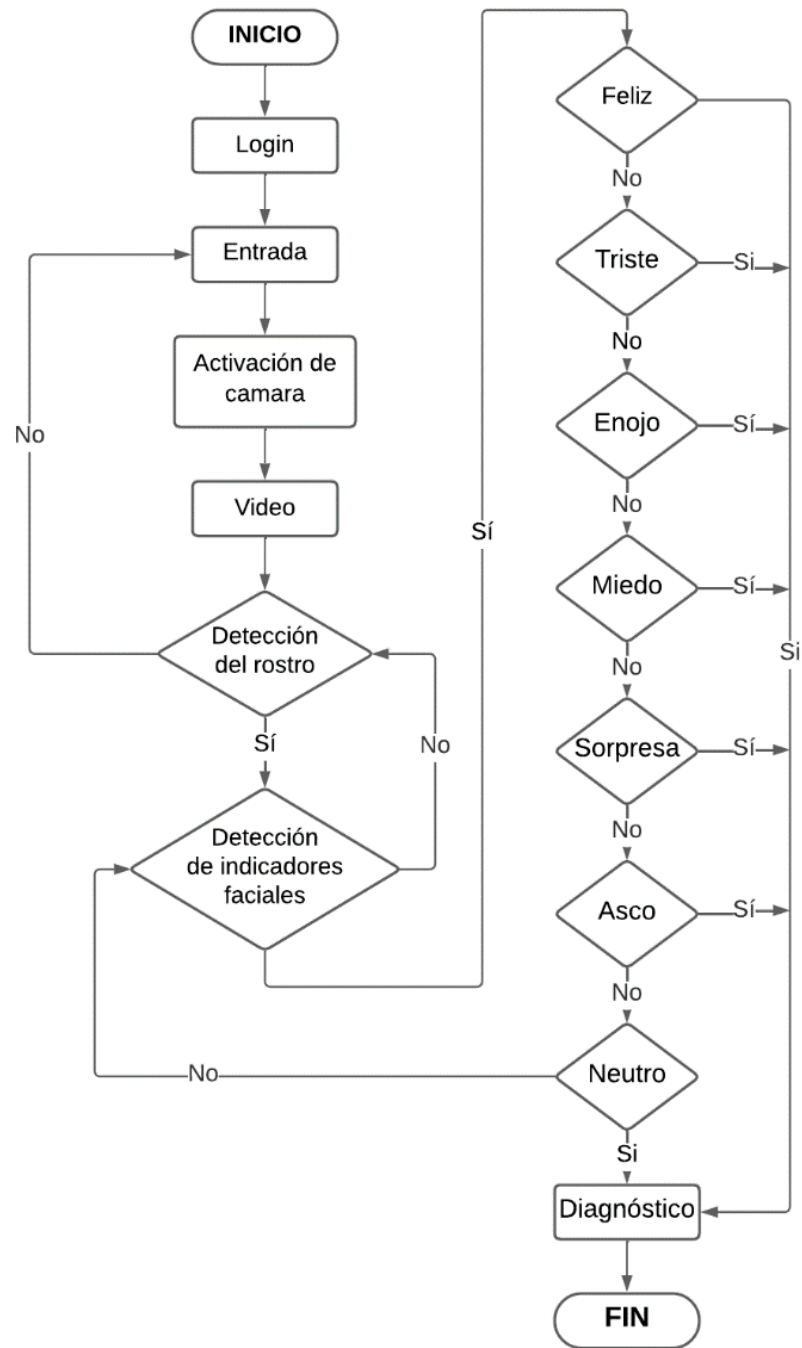


Figura 26. Diagrama de bloques del sistema

### 3.2.4 Diseño del sistema

#### a. Bibliotecas requeridas

Las librerías que se ocupó en este sistema se las detalla en la Tabla 28, que se presenta a continuación:



Tabla 28. Bibliotecas requeridas del sistema

Bibliotecas	Características
<b>OpenCV</b>	Se utiliza en la visión por computadora, se utiliza la librería cv2 para diferentes ámbitos como la captura de video, detección de rostros y manipulación de imágenes.
<b>Tensorflow.keras</b>	TensorFlow representa un entorno de aprendizaje automático de código abierto, un complemento importante es Keras que opera como una interfaz de nivel superior diseñada para la construcción y entrenamiento de modelos dentro de TensorFlow.
<b>NumPy</b>	Es una librería diseñada para el lenguaje de programación Python, la cual ayuda a la manipulación de matrices y ejecución de operaciones matemáticas. Su aplicación se centra en llevar a cabo diversas operaciones con datos de imágenes.
<b>Imutils</b>	Se trata de una biblioteca que ofrece funciones diseñadas para simplificar operaciones frecuentes en OpenCV, y su objetivo primordial es que el cambio de tamaño de una imagen sea eficiente y uniforme.
<b>Time</b>	Suministra funciones orientadas al manejo del tiempo. En la implementación presente, se emplea para evaluar el rendimiento del sistema y calcular los frames o cuadros por segundo.
<b>Tkinter</b>	Representa una (GUI) que no es más que una biblioteca gráfica de interfaz de usuario utilizada para el lenguaje de programación Python. En la presente implementación, se emplea para la creación de una interfaz gráfica que exhibe los parámetros y resultados del modelo y permite la visualización en tiempo real de video
<b>PIL (Python Imaging Library)</b>	Actualmente denominada Pillow, representa una librería de procesamiento digital de imágenes que se centra en la manipulación y visualización de imágenes en la interfaz gráfica.
<b>API</b>	hace referencia a un conjunto de funciones o métodos destinados a la interacción con servidores externos o APIs.
<b>Datetime</b>	Son clases orientadas al manejo de fechas y horas, se aplica para el registro temporal de acciones específicas en fechas y horas actualizadas en la interfaz gráfica.

- ***Inclusión de bibliotecas en la programación***

```
import cv2
from tensorflow.keras.applications.imagenet_utils import
preprocess_input
from tensorflow.keras.preprocessing.image import img_to_array
from tensorflow.keras.models import load_model
import numpy as np
import imutils
import time
import tkinter as ctk
import tkinter.messagebox as tkmb
from PIL import Image
from PIL import ImageTk
```

```
from api import *
from tkinter import ttk
from datetime import datetime
```

### ***b. Adquisición de la imagen***

Para que el sistema capte la imagen que será procesada del video en tiempo real se sigue una serie de pasos que se muestran en la Figura 27. Diagrama de flujo de la etapa de adquisición Figura 27:

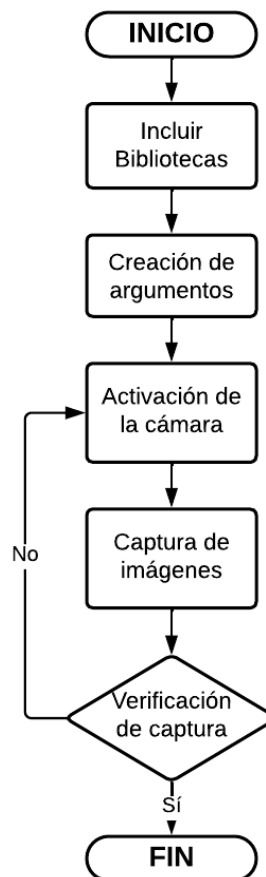


Figura 27. Diagrama de flujo de la etapa de adquisición.

La obtención de imágenes se lleva a cabo mediante el empleo de la cámara IMX219. Este diagrama describe de manera secuencial cada fase del procedimiento de adquisición, desde la inclusión de librerías, activación de la cámara hasta la obtención y almacenamiento de la imagen capturada, la cual, deberá contar con una calidad alta para su posterior procesamiento y entrada en el sistema de reconocimiento de emociones.

- **Inclusión de bibliotecas.** Se la utiliza la biblioteca de cv2 de OpenCV para la adquirir la visión mediante el computador.

```
import cv2
```

- **Creación de argumentos.** La adquisición de frames en este sistema se lo realizó mediante una función llamada "visualizar", la cual utiliza OpenCV y se almacenara toda la etapa de adquisición, adicional a esto, se estableció la creación de un objeto llamado "cap" para capturar imágenes desde una cámara web en tiempo real.

- **Activación de la cámara.** Inicializamos la cámara y le asignamos la tare de capturar frames a través del video en tiempo real

```
cap=cv2.VideoCapture(0,cv2.CAP_GSTREAMER)
```

Esta parte del código indica que se utilizará la cámara web conectada y se ocupara GStreamer como framework para la captura de video cuyo objetivo es obtener los frames a partir de interacciones con el hardware de la cámara. GStreamer se basa en un conjunto de librerías y herramientas que manejan y procesan datos, estos pueden ser de capturas de imágenes, reproducción, transmisión y manipulación de video y audio.

- **Captura de imágenes .** Captura de un solo cuadro o frames del video en tiempo real por medio de la variable "frame".

```
frame=cap.read()
```

- **Verificación de la captura exitosa.** Mediante la sentencia if decidimos las condiciones de que la variable "ret" indica que si sigue a la siguiente etapa de procesamiento y i no fuera el caso se liberara los valores de "ret" y regresar a la condición.

```
if cap is not None:
    ret, frame=cap.read()
else:
    cap.release()
```

### c. *Procesamiento*

Para que el sistema procese la imagen que será entrada al modelo de redes neuronales se sigue una serie de pasos que se muestra en la Figura 28.

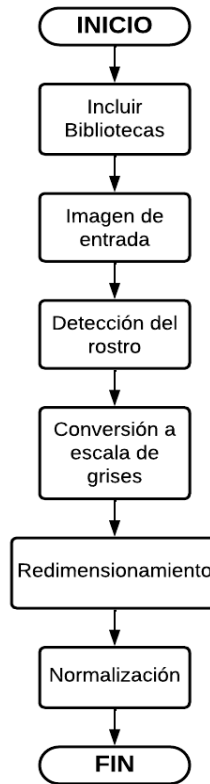


Figura 28. Diagrama de flujo del procesamiento de imágenes.

- ***Inclusión de bibliotecas.*** Se debe establecer las librerías y bibliotecas correctas para que la imagen sea procesada de manera eficiente, debido a que la próxima etapa es la entrada al modelo de redes neuronales.

```
from tensorflow.keras.applications.imagenet_utils import  
preprocess_input  
from tensorflow.keras.preprocessing.image import img_to_array  
from tensorflow.keras.models import load_model
```

- ***Frames de entrada.*** En esta etapa ya se verifico que la imagen esta correctamente obtenida y lista para su procesamiento
- ***Detección del rostro.*** Para el tratamiento de la imagen se debe enfocar en los rostros de las personas, para esto, se crea un blob de la imagen de entrada “frame” para la utilización del modelo preentrenado FaceNet que se carga con las bibliotecas de

OpenCV con el fin de extraer las regiones de interés de la imagen para optimizar los datos del frame antes de la entrada al modelo de redes neuronales.

```
blob = cv2.dnn.blobFromImage(frame, 1.0, (224, 224), (104.0, 177.0, 123.0))
```

- **Conversión a escala de grises.** La imagen del rostro se convierte a escala de grises para reducir su dimensionalidad y evitar las variaciones de color irrelevantes, esto se realizó mediante el siguiente comando:

```
face = cv2.cvtColor(cv2.imread(imaget_path), cv2.COLOR_BGR2GRAY)
```

- **Redimensionamiento.** Después de que la imagen se encuentra en escala de grises es crucial redimensionarla a un tamaño de 48x48 pixeles, debido a que el modelo de reconocimiento facial debe tener imágenes de este tamaño en su entrada, esto se realizó con el siguiente comando:

```
face = cv2.resize(face, (48, 48))
```

- **Normalización.** En este proceso la imagen que se encuentra en escala de grises y redimensionada se debe convertir a un array de NumPy, el siguiente comando convierte la imagen que se presenta como un objeto de PIL a un arreglo matemático NumPy.

```
face2 = img_to_array(face)
```

A continuación, es factible añadir una dimensión extra al array, esto con el fin de agregar una dimensión al lote ya que el modelo espera en su entrada un lote de imágenes

```
face2 = np.expand_dims(face2,axis=0)
```

#### **d. Modelo de inteligencia artificial**

- **Modelo Secuencial.** Se trata de un modelo de tipo secuencial, donde su estructura sigue la progresión lógica de capas desde la entrada hasta la salida. Este modelo demuestra un rendimiento óptimo en tareas de clasificación de imágenes y muestra compatibilidad con diversas herramientas dentro del ámbito del aprendizaje profundo.

La elección de este modelo también consideró los parámetros necesarios y la complejidad de nuestro sistema de detección de emociones.

- **Modelo también convolucional con varias capas.** El modelo adoptará una estructura multicapa, ya que sigue la arquitectura de una red neuronal, incluyendo capas de entrada, capas ocultas y capa de salida, donde diversas neuronas estarán interconectadas entre sí.
- **Aprendizaje profundo.** La elección de este aprendizaje se justifica por la construcción y entrenamiento de redes neuronales para realizar tareas de clasificación, el modelo multicapa escogido brinda la capacidad extraer y aprender representaciones de los datos, este tipo de aprendizaje se ha destacado en campos como la visión artificial, reconocimiento e patones, procesamiento de imágenes, procesamiento del lenguaje natural, juegos, etc.
- **Instalación de dependencias.** Con el siguiente comando se instala la biblioteca de TensorFlow y los resultados se presentan en la Figura 29:

```
!pip install tensorflow==2.14.0
Requirement already satisfied: gast!=0.5.0,!0.5.1,!0.5.2,>=0.2.1 in /usr/local/lib/python3.10/dist-packages (fr
Requirement already satisfied: google-pasta>=0.1.1 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2
Requirement already satisfied: h5py>=2.9.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14.0) (:
Requirement already satisfied: libclang>=13.0.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14
Requirement already satisfied: ml-dtypes==0.2.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14
Requirement already satisfied: numpy>=1.23.5 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14.0)
Requirement already satisfied: opt-einsum>=2.3.2 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.1
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14.0) (23
Requirement already satisfied: protobuf!=4.21.0,!4.21.1,!4.21.2,!4.21.3,!4.21.4,!4.21.5,<5.0.0dev,>=3.20.3 in
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14.0) (6
Requirement already satisfied: six>=1.12.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14.0) (:
Requirement already satisfied: termcolor>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2.14
Requirement already satisfied: typing-extensions>=3.6.6 in /usr/local/lib/python3.10/dist-packages (from tensorfl
Requirement already satisfied: wrapt<1.15,>=1.11.0 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /usr/local/lib/python3.10/dist-packages (fi
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /usr/local/lib/python3.10/dist-packages (from tensorflow==2
Collecting tensorboard<2.15,>=2.14 (from tensorflow==2.14.0)
  Downloading tensorboard-2.14.1-py3-none-any.whl (5.5 MB)
     |-----| 5.5/5.5 MB 98.7 MB/s eta 0:00:00
Collecting tensorflow-estimator<2.15,>=2.14.0 (from tensorflow==2.14.0)
  Downloading tensorflow_estimator-2.14.0-py2.py3-none-any.whl (440 kB)
     |-----| 440.7/440.7 kB 43.8 MB/s eta 0:00:00
```

Figura 29. Resultados de la instalación de la biblioteca TensorFlow.

Para la implementación del modelo de redes neuronales convolucionales basada en la arquitectura de ResNet se realizaron dos programaciones, una se realizó con 23 épocas y con 50 épocas. Rotar imágenes en un data set antes del entrenamiento es muy beneficioso debido que es una técnica de aumento de datos que ayuda a que el modelo

sea más versátil, al aumentar imágenes de diferentes perspectivas el modelo aprenderá a reconocer patrones característicos en diferentes perspectivas, aumentando su capacidad y evitando el sobreajuste

- **Data set**

En la búsqueda de una base de datos adecuada para el entrenamiento en reconocimiento de emociones, se priorizó la selección de conjuntos de datos que proporcionaran una amplia variedad de imágenes asociadas con emociones definidas. Este enfoque se adoptó con el propósito de mejorar la capacidad del modelo para generalizar y reconocer diversas expresiones emocionales. La estructura de las bases de datos adoptadas se organiza de manera específica como se observa en la Figura 30 para respaldar las fases críticas del proceso de entrenamiento.

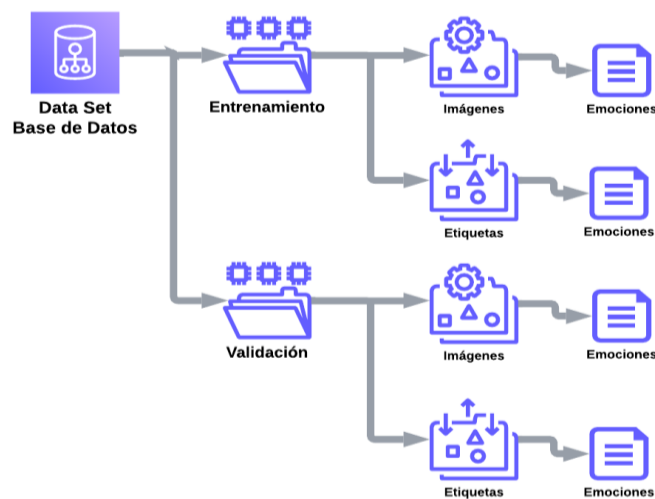


Figura 30. Arquitectura de la base de datos.

La base de datos se compone de dos componentes principales, ambas destinadas al proceso de entrenamiento. La primera componente es la carpeta de entrenamiento, que se desglosa en dos subcarpetas esenciales: "imágenes" y "etiquetas". En la subcarpeta de "imágenes", se almacenan imágenes representativas de cada emoción. En contraste, la subcarpeta de "etiquetas" contiene parámetros que caracterizan cada emoción, específicamente los indicadores faciales asociados.

La segunda componente corresponde a la carpeta de validación o pruebas del entrenamiento. Esta carpeta también se estructura en dos subcarpetas: "imágenes" y

"etiquetas". A diferencia de la carpeta de entrenamiento, las imágenes aquí presentes son desconocidas para la red neuronal durante el proceso de entrenamiento. Esta distinción es fundamental, ya que estas imágenes se utilizan para evaluar el rendimiento del modelo ante datos que no ha visto previamente. En consecuencia, esta fase de validación permite establecer un indicador crucial del nivel de desempeño del modelo, asegurando una evaluación rigurosa de su capacidad para generalizar y reconocer emociones de manera efectiva.

- **Características de la base de datos.** En el proceso de entrenamiento de la red neuronal convolucional, se ha logrado identificar y emplear un conjunto de datos específico con atributos definidos. El conjunto de datos seleccionado es conocido como "Facial Expression Recognition 2013 Dataset". Este conjunto de datos exhibe las siguientes características fundamentales:

- **Número de Imágenes:** Se compone de un total de 35,330 imágenes destinadas al reconocimiento de expresiones faciales.
- **Dimensionalidad:** Las imágenes se presentan en una dimensión, lo que implica un formato unidimensional para cada imagen en la base de datos.
- **Formato de Imágenes:** Las imágenes en este conjunto de datos se presentan en escala de grises (blanco y negro), proporcionando una representación visual sin la influencia de colores.
- **Emociones Por Detectar:** La base de datos aborda siete emociones distintas, que son ira, asco, miedo, feliz, triste, sorpresa y neutral. Estas emociones representan el conjunto de expresiones faciales objetivo para el modelo de reconocimiento.

Este conjunto de datos se seleccionó estratégicamente debido a sus dimensiones, formato de imágenes unidimensional y la diversidad de emociones contempladas, lo que lo hace adecuado para el entrenamiento de modelos de reconocimiento de expresiones faciales en el ámbito de la inteligencia artificial, según el libro [33] el porcentaje de entrenamiento recomendado para aplicaciones de visión artificial y para



base de datos no muy grandes es el de 80% y 20%, es el entrenamiento conveniente que se va a utilizar.

- **Descompresión de la data set.** En la Figura 31 se observa el proceso de descomprimir el data set:

```
!unrar x -Y "/content/Face_Emotion/images.rar" "/content"
```

```
Extracting /content/images/validation/happy/10468.jpg OK
Extracting /content/images/validation/happy/10480.jpg OK
Extracting /content/images/validation/happy/10528.jpg OK
Extracting /content/images/validation/happy/10540.jpg OK
Extracting /content/images/validation/happy/10552.jpg OK
Extracting /content/images/validation/happy/1056.jpg OK
Extracting /content/images/validation/happy/10571.jpg OK
Extracting /content/images/validation/happy/1058.jpg OK
Extracting /content/images/validation/happy/10622.jpg OK
Extracting /content/images/validation/happy/10638.jpg OK
Extracting /content/images/validation/happy/10640.jpg OK
Extracting /content/images/validation/happy/10644.jpg OK
Extracting /content/images/validation/happy/10647.jpg OK
```

Figura 31. Proceso de la descompresión del data set.

- **Resultados de la extracción.** A continuación, en la Figura 32 se observa los resultados de la extracción del data set

```
└─ images
  └─ train
    ├── angry
    ├── disgust
    ├── fear
    ├── happy
    ├── neutral
    ├── sad
    └── surprise
  └─ validation
    ├── angry
    ├── disgust
    ├── fear
    ├── happy
    ├── neutral
    ├── sad
    └── surprise
```

Figura 32. Resultados de la extracción del data set.

- **Aumento de imágenes.** Este parámetro es utilizado para el entrenamiento con la finalidad de robustecer el modelo y sus comandos para añadir las imágenes se rotaron son:

```

if __name__ == "__main__":
    """ enviando """
    np.random.seed(42)

    """ cargando dataset """
    data_path = "/content/images/train"
    data_path_test = "/content/images/validation"

    (train_x), (test_x) = load_data(data_path)
    """ Data augmentation """
    augment_data(train_x, data_path, augment=True) #si es True no se
aplica argumentation
    augment_data(test_x, data_path_test, augment=False)

```

Los resultados de la rotación de la imágenes y aumento de datos en el data set, que automáticamente se guardan en las carpetas específicas de cada emoción se puede observar en la Figura 33:

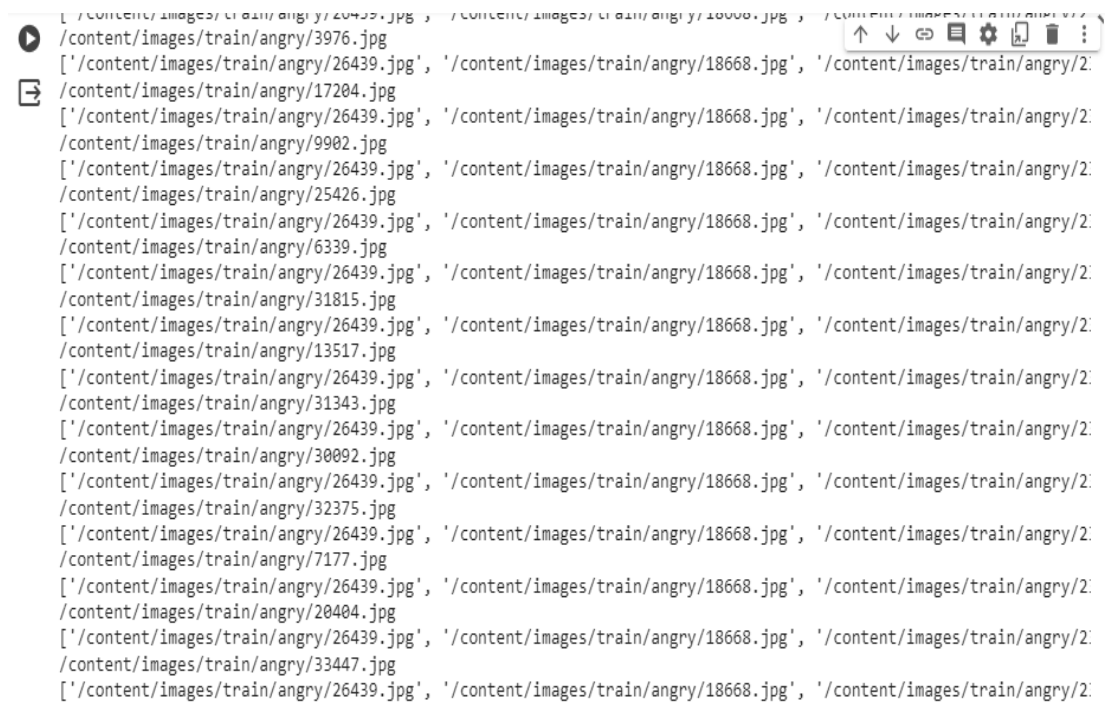


Figura 33. Resultados del aumento de datos.

- **Rutas del entrenamiento del data set**

```
train_data_dir = '/content/images/train'  
val_data_dir = '/content/images/validation'
```

- **Configuración del modelo.** Se definen parámetros importantes, se varía las épocas de 23 para el primero y 50 para el segundo modelo, para poder comparar y determinar el mejor modelo con menos errores y mejor precisión:

```
width_shape = 48  
height_shape = 48  
num_classes = 7  
epochs = 50  
batch_size = 32  
class_names =  
['angry', 'disgust', 'fear', 'happy', 'neutral', 'sad', 'surprise']
```

Se configura el dataset de entrenamiento y validación con los siguientes comandos:

```
train_datagen = ImageDataGenerator()  
val_datagen = ImageDataGenerator()  
  
train_generator = train_datagen.flow_from_directory(  
    train_data_dir,  
    target_size=(width_shape, height_shape),  
    batch_size=batch_size,  
  
    color_mode='grayscale',  
    class_mode='categorical', shuffle=True)  
val_generator = val_datagen.flow_from_directory(val_data_dir,  
    target_size=(width_shape, height_shape),  
  
    batch_size=batch_size,  
    color_mode='grayscale',  
    class_mode='categorical', shuffle=True)
```

Como se observa en la Figura 34 el número de imágenes para entrenamiento aumentó para las 7 clases o resultados esperados.

```
Found 50784 images belonging to 7 classes.  
Found 7066 images belonging to 7 classes.
```

Figura 34. Imágenes de entrenamiento y validación.

- **Arquitectura del Modelo.** Para determinar las capas del algoritmo, la red neuronal convolucional se utilizó la arquitectura ResNet, estos modelos emplean conexiones de atajo o saltos para eludir determinadas capas en las redes neuronales cuentan con capas que tienen convoluciones, no linealidades y la normalización por lotes.

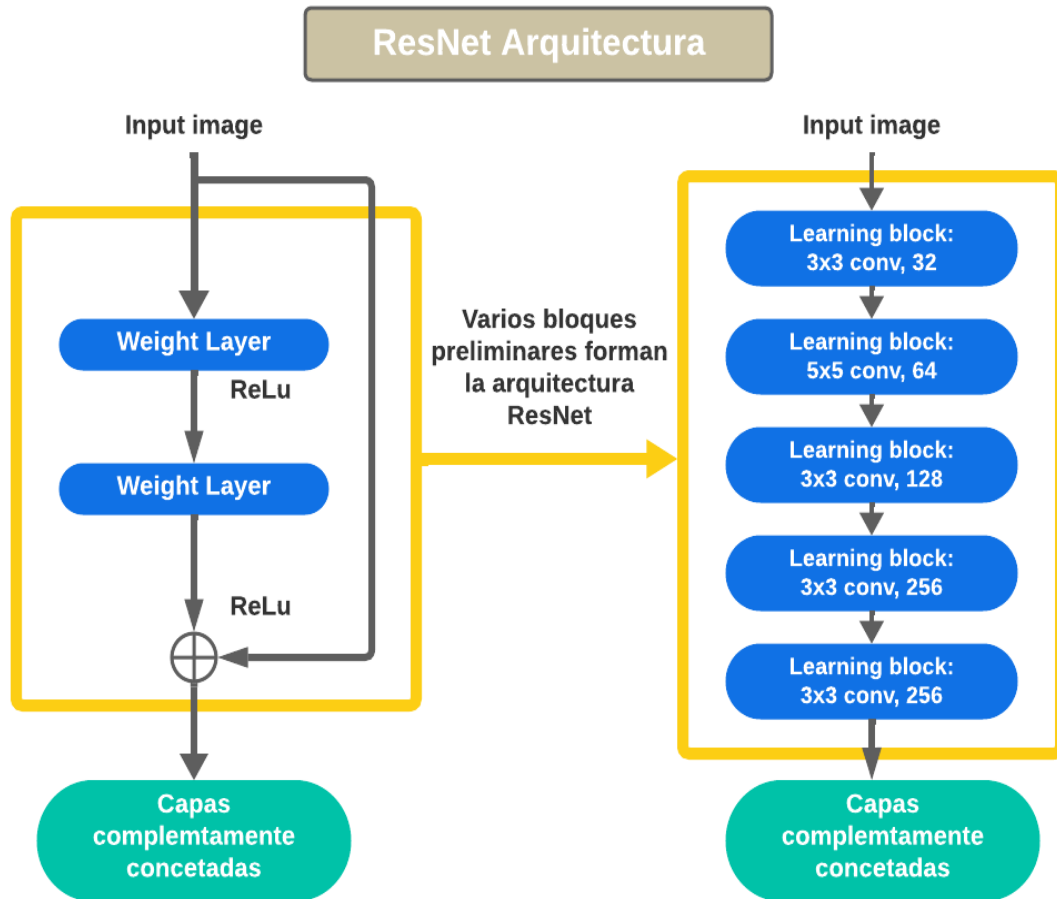


Figura 35.Arquitectura de ResNet

En base a la Figura 35 de la arquitectura, se establecen las primeras capas del modelo que son de convolución y Pooling, junto con las características de activación y normalización por lotes.

- **Capas del modelo.** En la Figura 36 se detallan las capas que tiene la red neuronal convolucional, comenzando desde la capa de entrada, capa oculta y capa de salida y aumentado la entrada que en este caso son los frames o imágenes detectados por la cámara web y de salida es la toma de decisión en la clasificación de las emociones definidas:

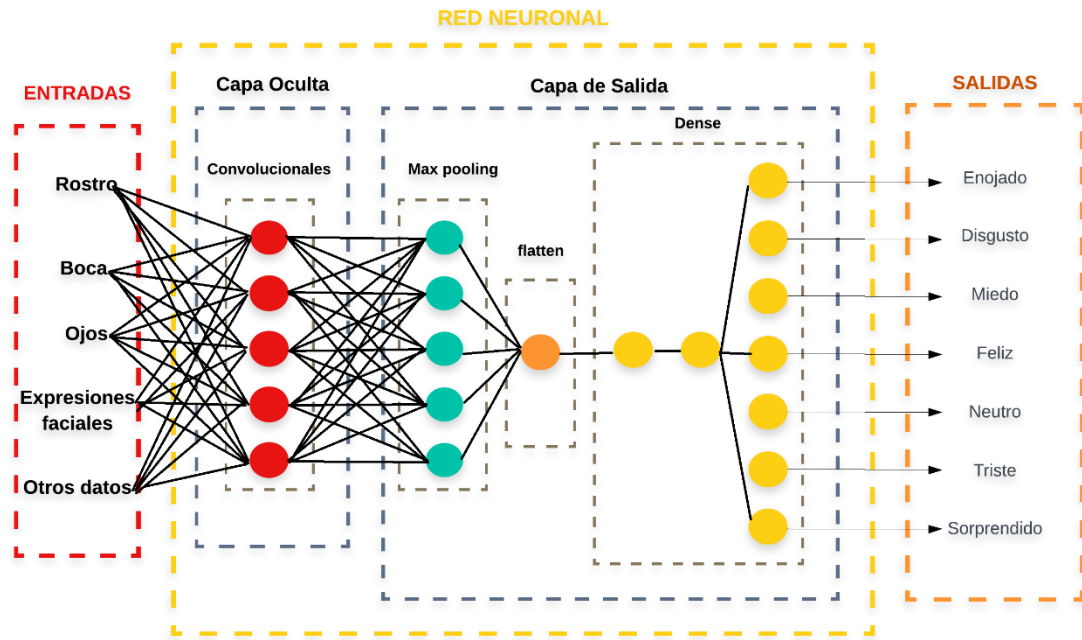


Figura 36. Capas del algoritmo

- **Capas de convolución.** Se utilizaron 6 de estas capas (Conv2D) que ayudan a extraer y analizar los patrones característicos o indicadores faciales de la imagen de entrada, para esto se debe aplicar filtros que se deslizan por cada núcleo o kernel obteniendo matrices de patrones y reduciendo la imagen con las características más importantes, a continuación de esta capa se aplica una función de normalización por lotes que consiste en uniformar las activaciones y así conseguir acelerar el entrenamiento.
- **Capas de Pooling.** Se utilizó las capas MaxPooling2D las cuales reducen las dimensiones espaciales de las imágenes después de pasar por convolución manteniendo los patrones más característicos.
- **Capas Flatten.** Se utilizó para convertir los datos de salida tridimensionales en un vector unidimensional para el manejo de operaciones matemáticas
- **Capas dense.** Son capas que se encuentran conectadas, se utilizan para realizar la clasificación a partir de las características extraídas, se utilizan funciones como ReLu y Dropout que se utilizan para estandarizar los datos de salida y darle un mejor rendimiento al modelo, por último, la activación softmax que transforma los resultados

de cada salida en probabilidades para la clasificación multicategórica y así para determinar los aciertos y errores

- **Definición del modelo neuronal.** Mediante los siguientes comandos se define el modelo neuronal requerido para los dos modelos:

```
from tensorflow.keras.models import Sequential, Model
from tensorflow.keras.layers import Conv2D, MaxPool2D, Dense, Flatten,
Dropout,
BatchNormalization, Input, AveragePooling2D, Activation
from tensorflow.keras.optimizers import Adam
model = Sequential()
```

Para la extracción de las características de las imágenes se decide establecer la siguiente estructura de comandos:

```
model.add(Conv2D(32, (3,3), padding='same',
input_shape=(width_shape,height_shape,1)))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPool2D(pool_size=(2,2)))
model.add(Dropout(0.2))
model.add(Conv2D(64, (5,5), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPool2D(pool_size=(2,2)))
model.add(Dropout(0.2))
model.add(Conv2D(128, (3,3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPool2D(pool_size=(2,2)))
model.add(Dropout(0.2))
model.add(Conv2D(256, (3,3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPool2D(pool_size=(2,2)))
model.add(Dropout(0.2))
model.add(Conv2D(512, (3,3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPool2D(pool_size=(2,2)))
model.add(Dropout(0.2))
```

Se clasificó de acuerdo con nuestras emociones ya definidas

```
model.add(Flatten())
model.add(Dense(128))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Dropout(0.2))
model.add(Dense(256))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Dropout(0.2))
model.add(Dense(num_classes, activation='softmax'))
```

- **Compilación del modelo.** Se revisa el modelo con el siguiente comando:

```
model.summary()
opt = Adam(learning_rate=1e-4)
model.compile(loss='categorical_crossentropy', optimizer=opt,
metrics=['accuracy'])
```

Se visualizó en la Figura 37 la compilación eficiente del modelo:

```
batch_normalization_22 (Ba (None, 256)          1024
tchNormalization)

activation_22 (Activation) (None, 256)          0

dropout_22 (Dropout)      (None, 256)          0

dense_8 (Dense)           (None, 7)            1799

=====
Total params: 1706759 (6.51 MB)
Trainable params: 1704007 (6.50 MB)
Non-trainable params: 2752 (10.75 KB)
```

Figura 37. Compilación del modelo de la red neuronal.

- **Código del entrenamiento.** Mediante estos códigos se procede a entrenar nuestro modelo

```
model.fit(
train_generator,
epochs=epochs,
validation_data=val_generator,
steps_per_epoch=train_generator.n//batch_size,
validation_steps=val_generator.n//batch_size,
callbacks=[tensorboard_callback])
```

- **Verificación de datos almacenados.** Al utilizar los datos generados por OpenCV para el reconocimiento facial se genera datos a partir de parámetros característicos del rostro, para esto se dividen en partes para determinar valores de estos puntos, como se observa en la Figura 38 se dividen los datos en cara, ojos, boca, expresiones faciales, mascara, movimiento de cabeza, perfil y a continuación los datos predictores.

```

▶ Data-pd.DataFrame()

#facedata

Data['PS1']-face1
Data['PS2']-face2
Data['PS3']-pface3
Data['PS4']-face4
Data['PS5']-faces
Data['PS6']-face6

# eyed data

Data['FS1']-eyed1
Data['FV2']-eyed2

#temperature data

Data['TS1']-mouth1
Data['TS2']-mouth2
Data['TS3']-mouth3
Data['TS4']-mouth4

#feelingsdata

Data['CE']-feelingsCE
Data['CP']-feelingsCP

#Other data

Data['EPS1']-mask_efficiency
Data['VS1']-movement_1
Data['SE']-factorfi

#prediction data

Data['PL']-pleak
Data['CD']-cCondition
Data['VC']-vCondition Data['HA']-Acc
Data['Stable']-stable

```

Figura 38. Datos almacenados de entrada.



- **Histograma de valores.** A partir de esto, se desarrolla funciones destinadas a solicitar los valores inherentes a cada elemento, permitiendo así determinar los pesos asociados a cada característica, como se observa en la Figura 39.

```

21    PS5    2205 non-null    float64
22    PS6    2205 non-null    float64
23    FS1    2205 non-null    float64
24    FV2    2205 non-null    float64
25    TS1    2205 non-null    float64
26    TS2    2205 non-null    float64
27    TS3    2205 non-null    float64
28    TS4    2205 non-null    float64
29    CE     2205 non-null    float64
30    CP     2205 non-null    float64
31    EPS1   2205 non-null    float64
32    VS1   2205 non-null    float64
33    SE     2205 non-null    float64

Dtypes: float(17), int64(5)
Memory usage: 379.1 KB

```

Figura 39. Guardado de datos.

- **Distribución normal.** Utilizando la información recopilada en la figura anterior, se genera gráficos de distribución normal, como se visualiza en la Figura 40 ninguno de los conjuntos sigue una distribución normal esto significa que cada conjunto de datos se adapta para que sus parámetros se estandaricen con los valores presentes en la base de datos.

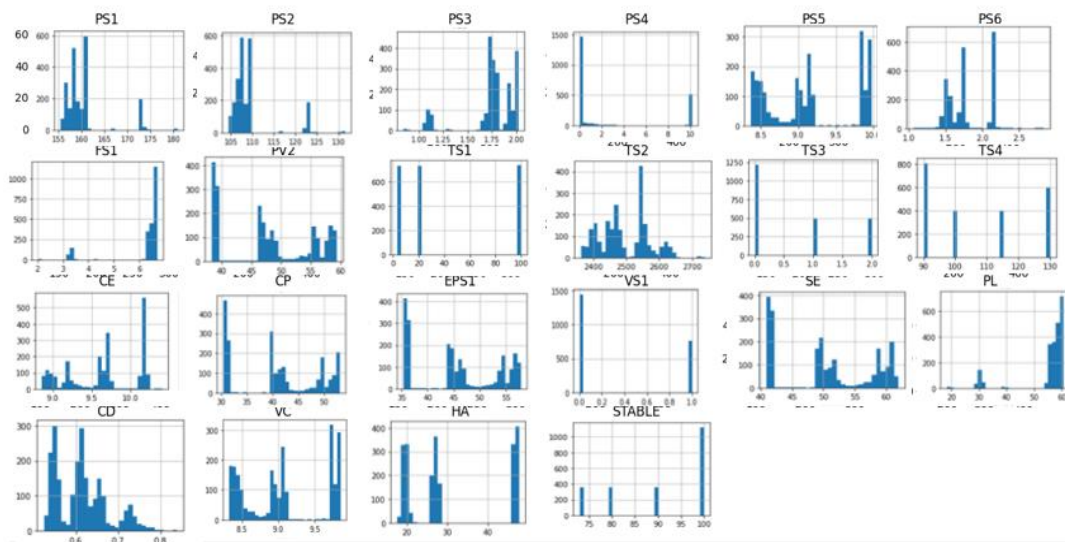


Figura 40. Distribución normal.

- **Matriz de correlación.** A partir de los valores graficados en la distribución normal, se llevó a cabo la construcción de una matriz de correlación presente en la Figura 41 entre los conjuntos de datos de entrada, abarcando tanto el total de valores de entrenamiento como la totalidad de los valores asociados a cada una de las variables, la función principal es medir la relación lineal entre cada variable y sus valores pasados.



Figura 41. Matriz de correlación.

- **Modelo generado.** Una vez creada la matriz de correlación, el siguiente paso es analizar la misma para poder establecer los pesos para el modelo de ganancia de información, con la compilación del modelo que se observa en la Figura 42 se despliegan los pesos de cada iteración, a partir de esto se realizó las gráficas de los datos que se presentan en la Figura 43.

```
model.compile(loss='categorical_crossentropy')  
  
PS1      0.580035  
PS2      0.591373  
PS3      0.140037  
PS4      0.088636  
PS5      0.353748  
PS6      0.354115  
FS1      0.371377  
FV2      0.252797  
TS1      0.436819  
TS2      0.406083  
TS3      0.439246  
TS4      0.376826  
EC       0.129169  
CP       0.062581  
EPS1     0.254181  
VS1     0.120241  
HE       0.415622  
dtype: float64
```

Figura 42. Valores de los pesos.

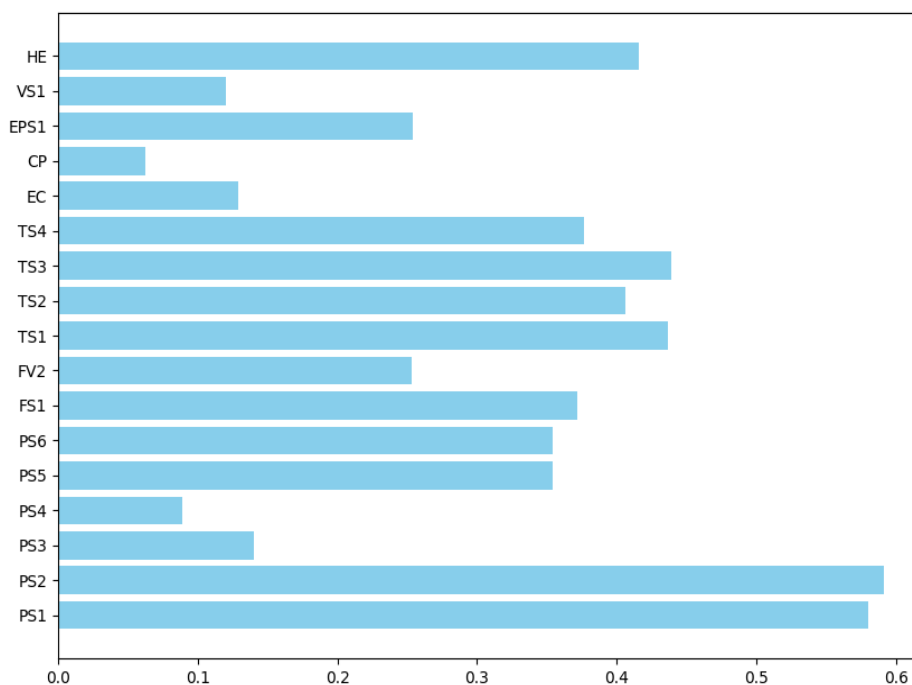


Figura 43. Representación gráfica de los pesos.

- **Resultados del entrenamiento.** Como se observa en la Figura 44 y Figura 45 a medida que las épocas suben el modelo va generando menos errores:

- **Modelo entrenado con 23 épocas**

```
1587/1587 [=====] - 634s 399ms/step - loss: 0.6625 - accuracy: 0.7521 - val_loss: 1.1184 - val_accuracy: 0.5844
Epoch 14/23
1587/1587 [=====] - 631s 398ms/step - loss: 0.6504 - accuracy: 0.7545 - val_loss: 1.0730 - val_accuracy: 0.6003
Epoch 15/23
1587/1587 [=====] - 628s 396ms/step - loss: 0.6407 - accuracy: 0.7607 - val_loss: 1.0787 - val_accuracy: 0.5980
Epoch 16/23
1587/1587 [=====] - 607s 382ms/step - loss: 0.6264 - accuracy: 0.7670 - val_loss: 1.0551 - val_accuracy: 0.6074
Epoch 17/23
1587/1587 [=====] - 606s 382ms/step - loss: 0.6206 - accuracy: 0.7697 - val_loss: 1.0604 - val_accuracy: 0.6026
Epoch 18/23
1587/1587 [=====] - 612s 386ms/step - loss: 0.6072 - accuracy: 0.7712 - val_loss: 1.0611 - val_accuracy: 0.6021
Epoch 19/23
1587/1587 [=====] - 606s 382ms/step - loss: 0.5939 - accuracy: 0.7774 - val_loss: 1.0705 - val_accuracy: 0.5993
Epoch 20/23
1587/1587 [=====] - 606s 382ms/step - loss: 0.5886 - accuracy: 0.7799 - val_loss: 1.1116 - val_accuracy: 0.6109
Epoch 21/23
1587/1587 [=====] - 607s 382ms/step - loss: 0.5755 - accuracy: 0.7868 - val_loss: 1.0345 - val_accuracy: 0.6156
Epoch 22/23
1587/1587 [=====] - 608s 383ms/step - loss: 0.5650 - accuracy: 0.7882 - val_loss: 1.0220 - val_accuracy: 0.6260
Epoch 23/23
1587/1587 [=====] - 610s 384ms/step - loss: 0.5593 - accuracy: 0.7913 - val_loss: 1.0359 - val_accuracy: 0.6274
<keras.src.callbacks.History at 0x79918c15b310>
```

Figura 44. Resultados del entrenamiento con 23 épocas.

- **Modelo entrenado con 50 épocas**

```
Epoch 41/50
900/900 [=====] - 19s 21ms/step - loss: 0.7953 - accuracy: 0.8227 - val_loss: 1.0288 - val_accuracy: 0.8103
Epoch 42/50
900/900 [=====] - 18s 19ms/step - loss: 0.7710 - accuracy: 0.8501 - val_loss: 1.0449 - val_accuracy: 0.8176
Epoch 43/50
900/900 [=====] - 19s 21ms/step - loss: 0.7611 - accuracy: 0.8883 - val_loss: 1.0371 - val_accuracy: 0.8745
Epoch 44/50
900/900 [=====] - 18s 20ms/step - loss: 0.7582 - accuracy: 0.8764 - val_loss: 1.0258 - val_accuracy: 0.8286
Epoch 45/50
900/900 [=====] - 18s 20ms/step - loss: 0.7350 - accuracy: 0.9005 - val_loss: 1.0456 - val_accuracy: 0.8666
Epoch 46/50
900/900 [=====] - 19s 21ms/step - loss: 0.7243 - accuracy: 0.8806 - val_loss: 1.0384 - val_accuracy: 0.8211
Epoch 47/50
900/900 [=====] - 18s 20ms/step - loss: 0.7088 - accuracy: 0.8673 - val_loss: 1.0562 - val_accuracy: 0.8496
Epoch 48/50
900/900 [=====] - 18s 20ms/step - loss: 0.6992 - accuracy: 0.8817 - val_loss: 1.0695 - val_accuracy: 0.8757
Epoch 49/50
900/900 [=====] - 18s 20ms/step - loss: 0.6867 - accuracy: 0.9136 - val_loss: 1.0351 - val_accuracy: 0.8987
Epoch 50/50
900/900 [=====] - 18s 20ms/step - loss: 0.6761 - accuracy: 0.9237 - val_loss: 1.0560 - val_accuracy: 0.9118
```

Figura 45. Resultados del entrenamiento con 50 épocas.

Google Colab proporciona con una interfaz gráfica llamada TensorBoard, es una herramienta que permite visualizar los datos del modelo de la red neuronal convolucional. Esta interfaz ofrece una manera intuitiva y poderosa de comprender los datos generados durante el entrenamiento, en la Figura 46 se observa los datos de entrenamiento y validación de la red neuronal convolucional a lo largo de las iteraciones del entrenamiento.

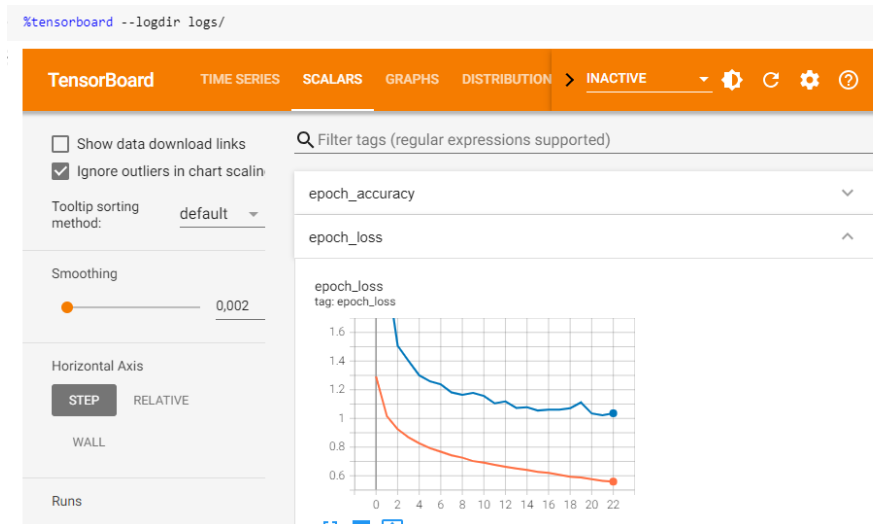


Figura 46. Entorno de TensorBoard.

- **Precisión de las épocas.** En el análisis de la precisión métrica a lo largo de múltiples épocas de entrenamiento, se observa un patrón ascendente tanto en la precisión del conjunto de entrenamiento como en la de validación. Este comportamiento sugiere un proceso de aprendizaje eficaz en el que el modelo ha logrado capturar de manera efectiva patrones inherentes a los conjuntos de datos de entrenamiento y validación. La convergencia simultánea de ambas métricas indica un equilibrio robusto en el proceso de aprendizaje, respaldando la capacidad del modelo para generalizar eficazmente a nuevos datos.

No obstante, al examinar los resultados representados en la Figura 47, se revela una discrepancia significativa, ya que la precisión en el conjunto de validación se encuentra considerablemente por debajo de la del conjunto de entrenamiento. Esta disparidad plantea interrogantes sobre la confiabilidad de los resultados, indicando la necesidad de una evaluación más detallada. Por otro lado, la Figura 48 presenta resultados preliminares prometedores, confirmando la validez y el progreso positivo del modelo durante el entrenamiento.

Es relevante destacar que, aunque la posibilidad de aumentar las épocas de entrenamiento existe, se observa que los cambios ya no serían bruscos ni representativos en términos de mejora en las métricas. Este hallazgo sugiere una etapa avanzada en el proceso de aprendizaje, donde las ganancias adicionales podrían ser

más graduales y se debe considerar cuidadosamente la relación costo-beneficio de continuar con un entrenamiento prolongado.

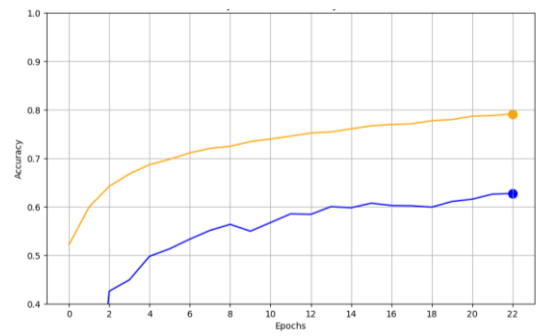


Figura 47. Precisión del modelo con 23 épocas.

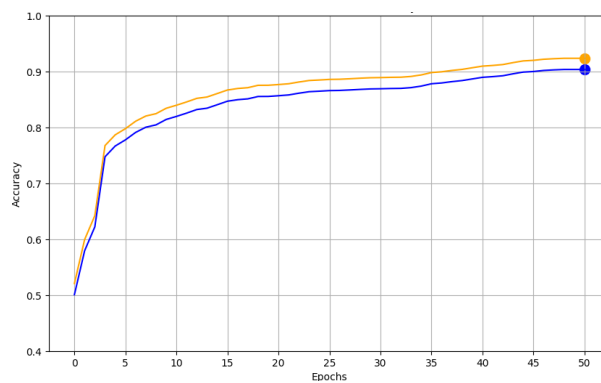


Figura 48. Precisión del modelo con 50 épocas.

- **Perdidas según las épocas.** La observación de la Figura 50 de pérdidas a lo largo de las épocas revela un patrón descendente tanto en las pérdidas de entrenamiento como en las de validación. Este comportamiento es indicativo de un progreso positivo en el entrenamiento del modelo. La disminución constante de las pérdidas de entrenamiento sugiere que el modelo está ajustándose eficazmente a los datos de entrenamiento, mientras que la disminución paralela de las pérdidas de validación indica que el modelo está generalizando bien a datos no vistos. La coherencia en la tendencia descendente de ambas curvas de pérdida respalda la convergencia exitosa del modelo hacia parámetros óptimos. Como se observa en la Figura 49 las pérdidas del entrenamiento disminuyen de manera favorable, mientras que la de validación no, en la Figura 50 los datos están bajando simultáneamente en las métricas de precisión, estos gráficos respaldan de manera positiva la efectividad y validez del modelo en el proceso de entrenamiento, incluso si se pretendiera aumentar épocas los cambios ya no serían bruscos ni representativos .

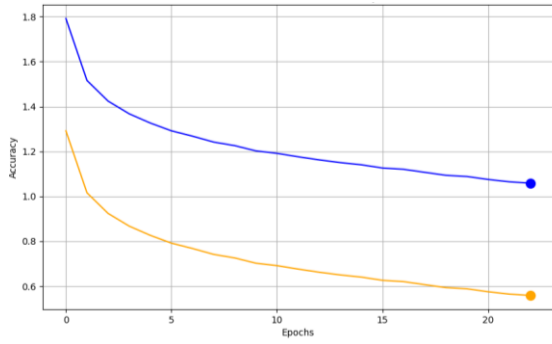


Figura 49. Pérdidas en el entrenamiento con 23 épocas.

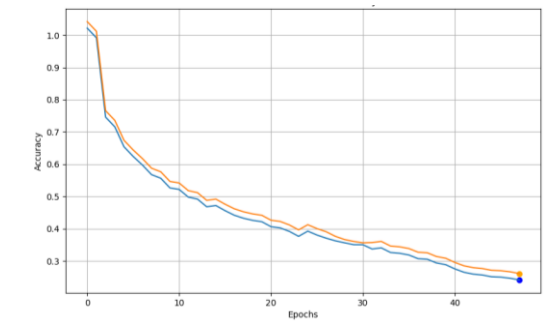


Figura 50. Pérdidas en el entrenamiento con 50 épocas.

- **Evaluación de resultados del entrenamiento**
- **Matriz de correlación.** A continuación, en la Figura 51 y Figura 52 se presentan las matrices de confusión de los resultados:

	1	2	3	4	5	6	7
1	521	16	76	54	114	154	25
2	17	69	8	3	2	8	4
3	108	3	439	46	101	240	81
4	46	2	36	1527	82	95	37
5	96	2	53	91	712	244	18
6	114	5	100	44	186	677	13
7	12	1	55	39	22	30	638
	1	2	3	4	5	6	7

Figura 51. Matriz de confusión del modelo con 23 épocas.

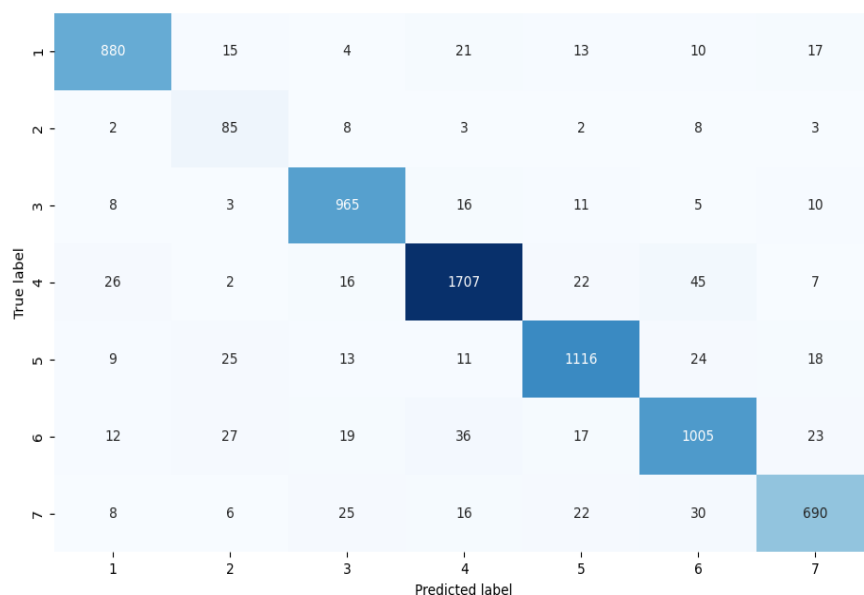


Figura 52. Matriz de confusión del modelo con 50 épocas.

En la Figura 53 y Figura 54 se visualiza los resultados del proceso de entrenamiento, abarcando los parámetros asociados a la validación del modelo de inteligencia artificial. Estas representaciones evidencian que el modelo entrenado a lo largo de 50 épocas presenta un rendimiento sustancialmente superior en comparación con el modelo entrenado durante 23 épocas. Este notable aumento en la eficacia se atribuye principalmente a la precisión del modelo, la cual se ve significativamente mejorada con el aumento en el número de épocas y es el utilizado en este sistema para una mejor detección de emoción.

Found 7066 images belonging to 7 classes				
	precision	recall	F1-score	Support
1	0.5700	0.5427	0.5560	960
2	0.7041	0.6216	0.6603	111
3	0.5724	0.4312	0.4919	1018
4	0.8465	0.8367	0.8416	1825
5	0.5841	0.5855	0.5848	1216
6	0.4675	0.5944	0.5234	1139
7	0.7819	0.8005	0.7911	797
Accuracy			0.6486	7066
macro avg	0.6466	0.6304	0.6356	7066
weighted	0.6537	0.6486	0.6484	7066

Figura 53. Resultados del modelo con 23 épocas



Found 7066 images belonging to 7 classes

	precision	recall	F1-score	Support
1	0.9312	0.9167	0.9239	960
2	0.5215	0.7658	0.6204	111
3	0.9190	0.9479	0.9333	1018
4	0.9431	0.9353	0.9392	1825
5	0.9277	0.9178	0.9227	1216
6	0.8917	0.8824	0.8870	1139
7	0.8984	0.8657	0.8818	797
Accuracy			0.9125	7066
macro avg	0.9095	0.8965	0.9017	7066
weighted	0.9357	0.9125	0.9123	7066

Figura 54. Resultados del modelo con 50 épocas

### e. Web hosting

Se estableció el alojamiento web en 000webhost, como se muestra en la Figura 55 es el panel principal, se creó un sitio web que ayuda como un entorno administrador de los datos recibidos y enviados y también para la creación de la interfaz administrativa para visualizar y gestionar el patrón de emociones.

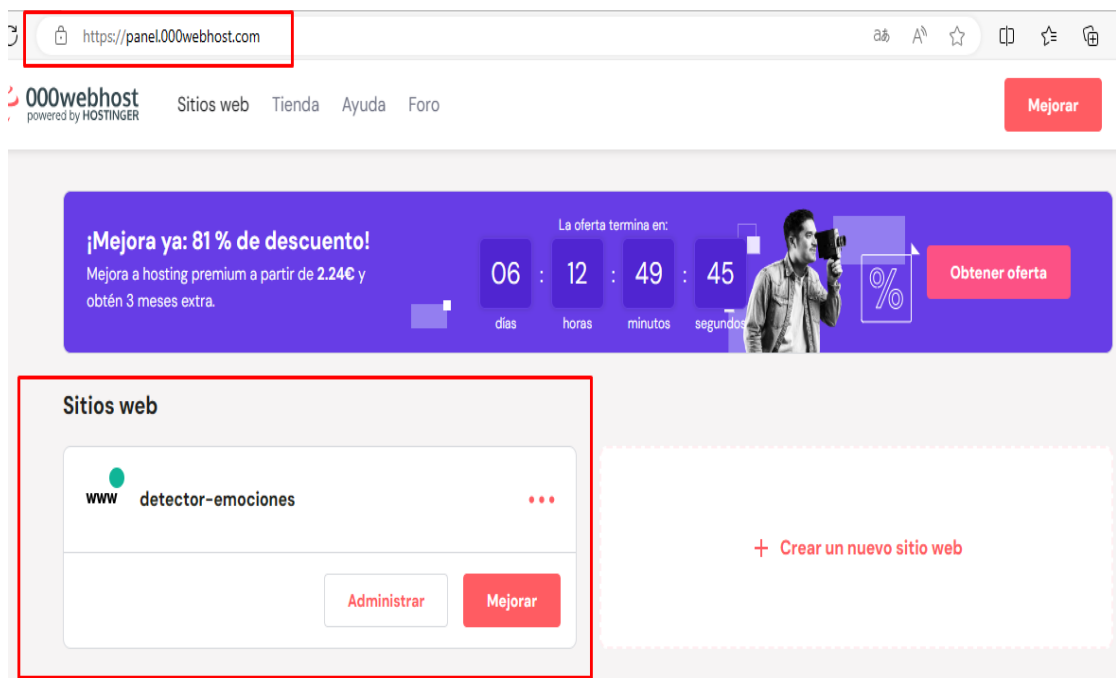


Figura 55. Ingreso al panel de 000web host.

Para almacenar los datos que se reciben desde el microrontrador Jetson Nano, se crea la base de datos en PhPMyAdmin como se observa en la Figura 56:



Figura 56. Administración de la base de datos.

Al entrar en el entorno de la base de datos de PhPMyAdmin, en la Figura 57 se puede observar las diferentes características y también el id de nuestra base de datos:

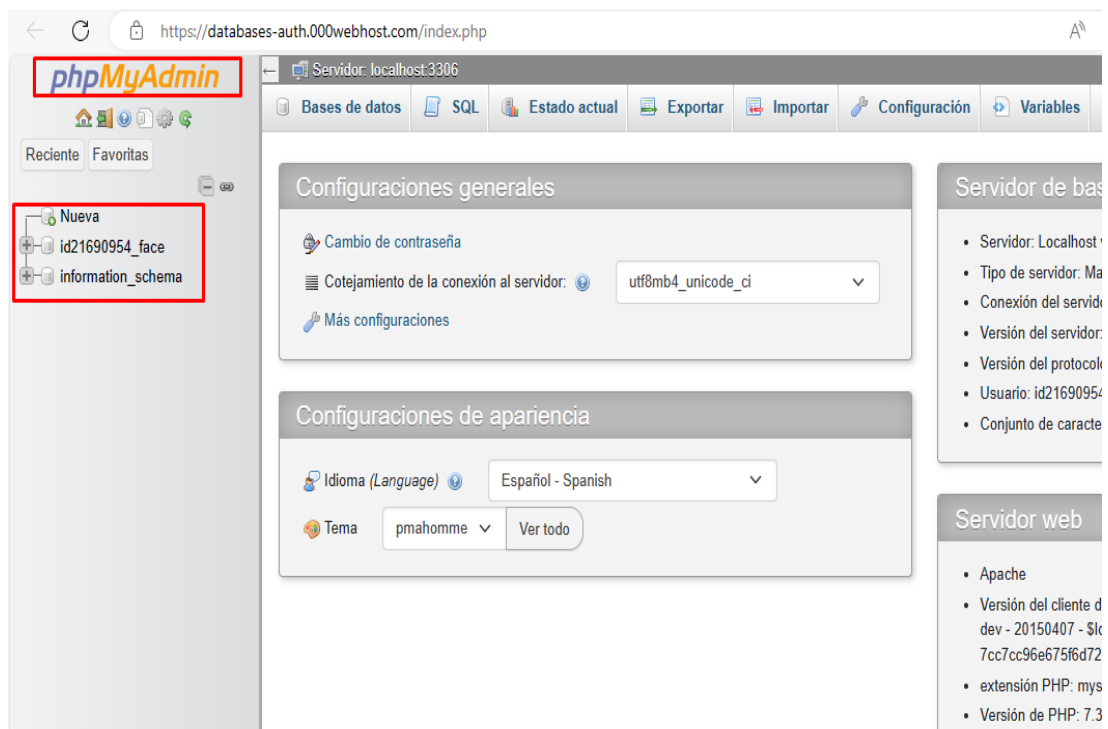


Figura 57. Entorno de base de datos.

A continuación, en la Figura 58 se indica la estructura y organización de las base de datos.

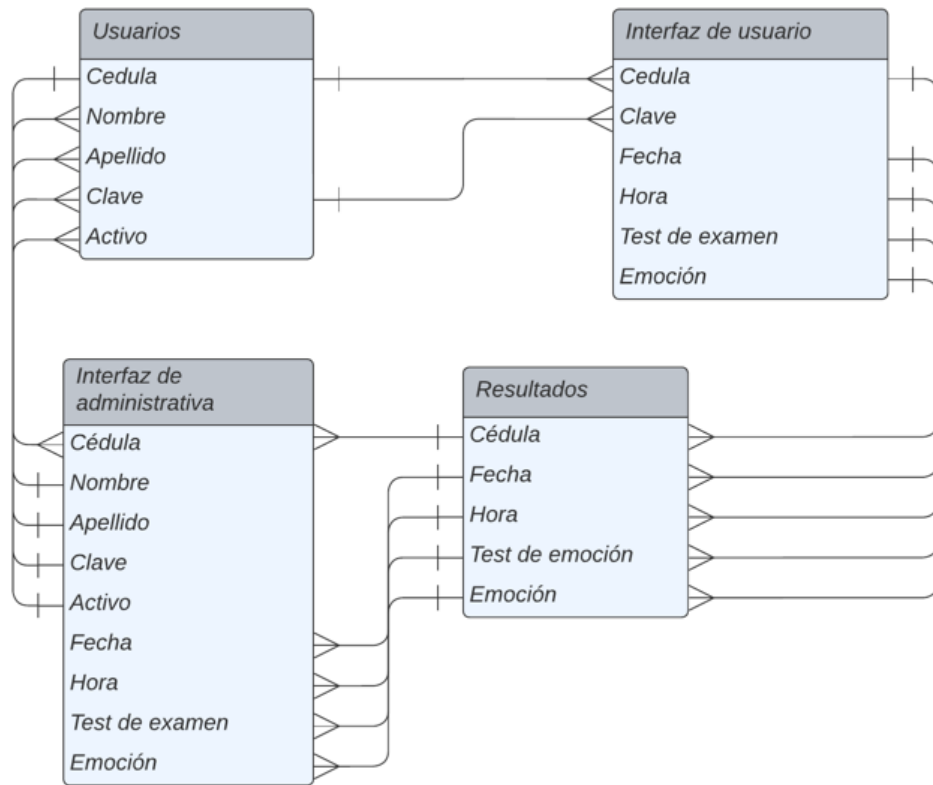


Figura 58. Representación de la base de datos

En el entorno de la base de datos, se visualiza los apartados creados para recibir los datos, en la Figura 59 se presenta todas las partes que tiene la base de datos:

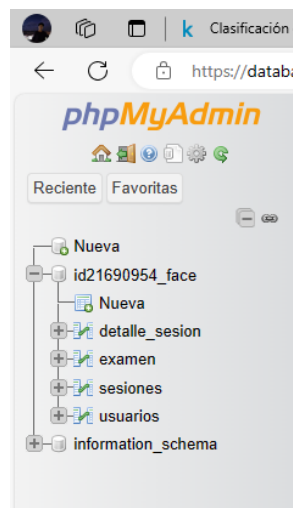


Figura 59. Apartados de la base de datos.

En la sección usuarios podemos observar los nombres de las personas que ayudaron para probar el sistema de detección de emociones, como se observan en la Figura 60:

Mostrando filas 0 - 15 (total de 16. La consulta tardó 0.0007 segundos.)

```
SELECT * FROM `usuarios`
```

Perfilando [ Editar en línea ] [ Editar ] [ Explicar SQL ] [ Crear código PHP ] [ Actualizar ]

Mostrar todo | Número de filas: 25 | Filtrar filas: Buscar en esta tabla | Ordenar según la clave: Ninguna

ID	CEDULA	NOMBRE	APELLIDO	ESTADO	CLAVE	TIPO_USUARIO
1	1804149233	marlon	Torres	a	12345	USUARIO
2	1804149234	Marlon	Maisanche	a	12345	USUARIO
3	1804420816	Daniel	Fiallos	a	12345	ADMINISTRADOR
10	1804420832	Fernando	Fiallos	a	12345	USUARIO
11	1803510864	Isabel	Araujo	a	12345	USUARIO
12	1804204988	David	Palan	a	12345	USUARIO
13	1850034057	Shirley	Pazmiño	a	12345	USUARIO
14	1803830668	Erick	Sandoval	a	12345	USUARIO
15	1850536242	Bryan	Viteri	a	12345	USUARIO
16	0202010047	Francisco	Moreno	a	12345	USUARIO
17	1722680269	Cristian	Chimbo	a	12345	USUARIO
18	1312721606	Dario	Garcia	a	12345	USUARIO

Figura 60. Base de datos de los usuarios

## f. Interfaces

- **Interfaz de usuario.** Se ha desarrollado una interfaz dedicada en la microcomputadora Jetson Nano con el propósito de simplificar la implementación del sistema propuesto, como se ilustra en la Figura 61. Con el objetivo de optimizar el reconocimiento de emociones, se ha implementado un proceso de registro y autenticación de usuarios en la interfaz. Una vez que el usuario ha iniciado sesión, se le brinda la opción de llevar a cabo la prueba de detección de emociones, seguida de la posibilidad de activar la cámara.

Es fundamental que, al encontrarse frente a la cámara, el usuario mantenga una distancia mínima de 30 cm para garantizar la calidad óptima del enfoque facial. Se recomienda además que los giros de la cabeza no excedan los 45°, asegurando así una determinación precisa de la emoción. Sin embargo, es importante destacar que la red neuronal implementada exhibe una elevada precisión, permitiendo la detección fiable de emociones incluso en rostros con inclinación o giros pronunciados.

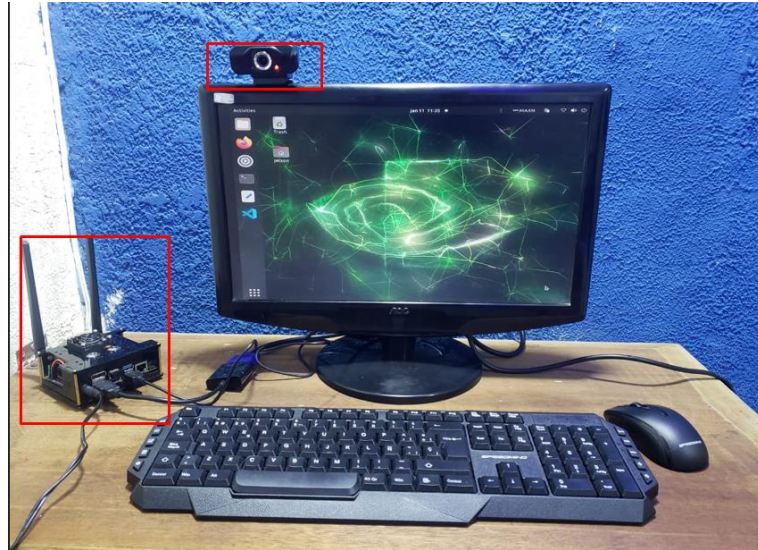


Figura 61. Posicionamiento del usuario con el sistema.

- **Acceso.** Se realizo la interfaz de usuario en Python con la herramienta tkinter debido al nivel optimizado que proporciona y se pueda ejecutar de manera eficiente en la Jetson Nano.

1. Para acceder a la interaz se debe colocar el usuario y contraseña antes registrados como se observa en la Figura 62:



Figura 62. Inicio de sesión en la interfaz de usuario.

2. En la Figura 63, se presenta diversas pruebas de emociones diseñadas específicamente para abordar diferentes contextos y necesidades. Estas pruebas incluyen evaluaciones psicológicas, exámenes académicos, evaluaciones laborales, mediciones de satisfacción al cliente, evaluaciones de seguridad, y una prueba de unidad. Los usuarios pueden seleccionar la prueba que mejor se adapte a sus necesidades y requisitos específicos.

The screenshot displays the 'Detector de emociones' interface. On the left, a user profile box shows the following information: 'Usuario', 'CEDULA: 1804149233', 'NOMBRE: marlon', 'APELLIDO: Torres', and 'ROL: USUARIO'. Below this is a green button labeled 'Cerrar Sesión'. The main area on the right lists four tests, each with a title, description, instructions, and a green 'EMPEZAR' button:

- Tema: Test de examen psicología**  
Descripción: Probar el desempeño de una persona al dar un prueba lógica matemática.  
Instrucciones: Encienda la cámara, para poder acceder a la prueba
- Tema: Test de examen académico**  
Descripción: Probar el desempeño de una persona al dar un prueba de español.  
Instrucciones: Encienda la cámara, para poder acceder a la prueba
- Tema: Test de examen laboral**  
Descripción: Probar el desempeño de una persona al dar un prueba de química.  
Instrucciones: Encienda la cámara, para poder acceder a la prueba
- Tema: Test de examen de satisfacción al cliente**  
Descripción: Probar la satisfacción del cliente  
Instrucciones: Encienda la cámara, para poder acceder a la prueba

Figura 63. Lista de test creados para determinar las emociones.

3. Como se indica en la Figura 64, al usuario le aparecerá otra ventana que indica iniciar el video a través de la cámara.

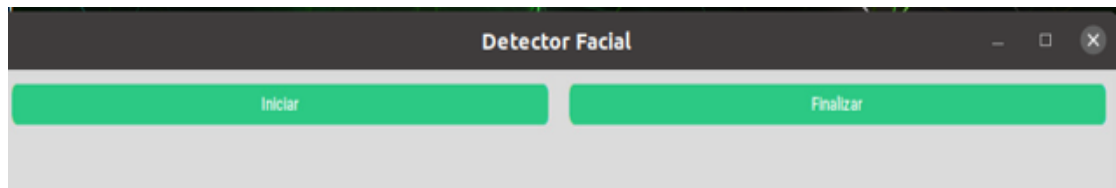


Figura 64. Interfaz de activación de cámara web.

4. Por último, se despliega la captación del frame del video en tiempo real conjuntamente con la detección de rostro y la emoción detectada como se observa en la Figura 65.

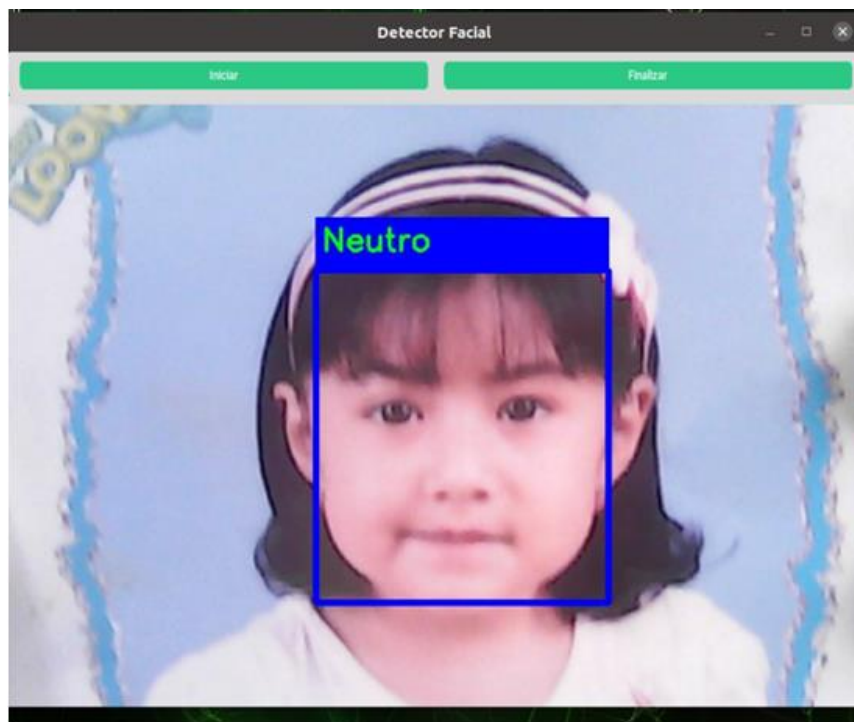


Figura 65. Activación del sistema.

- **Interfaz administrativa.** La interfaz que se observa en la Figura 66 se desarrolla utilizando el servicio de alojamiento web 000webhost. A través de esta plataforma, el usuario administrativo puede acceder para gestionar y visualizar todas las características del sistema. El acceso a la interfaz requiere credenciales de administrador, con un nombre de usuario y una clave proporcionados para autenticación.

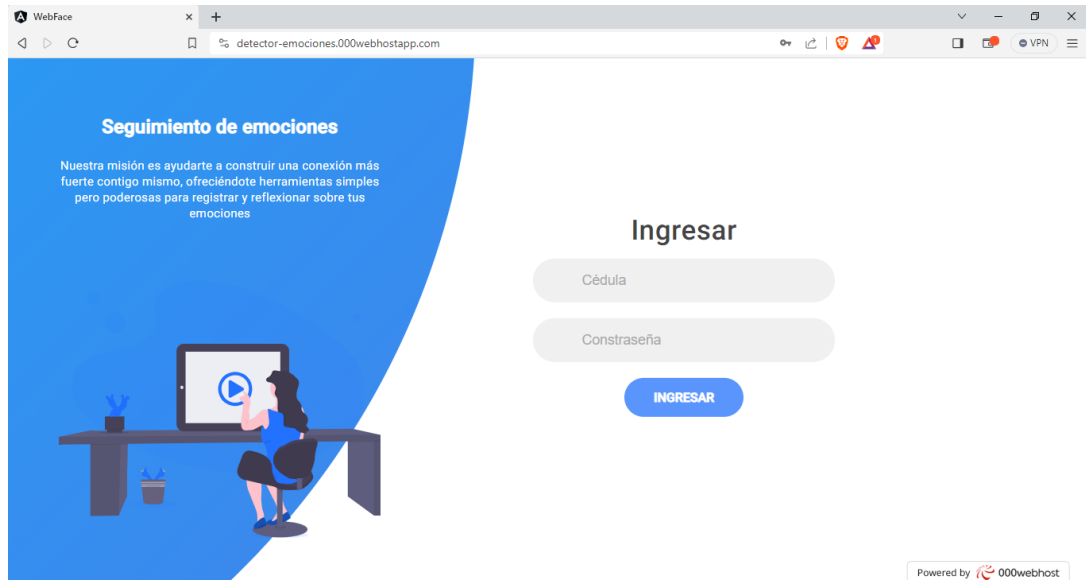


Figura 66. Ingreso a la interfaz administrativa.

Al acceder, el usuario administrativo se encontrará en la pestaña de exámenes, donde se puede verificar el test en uso, junto con la fecha correspondiente y el nombre de usuario asociado, como se muestra en la Figura 67.

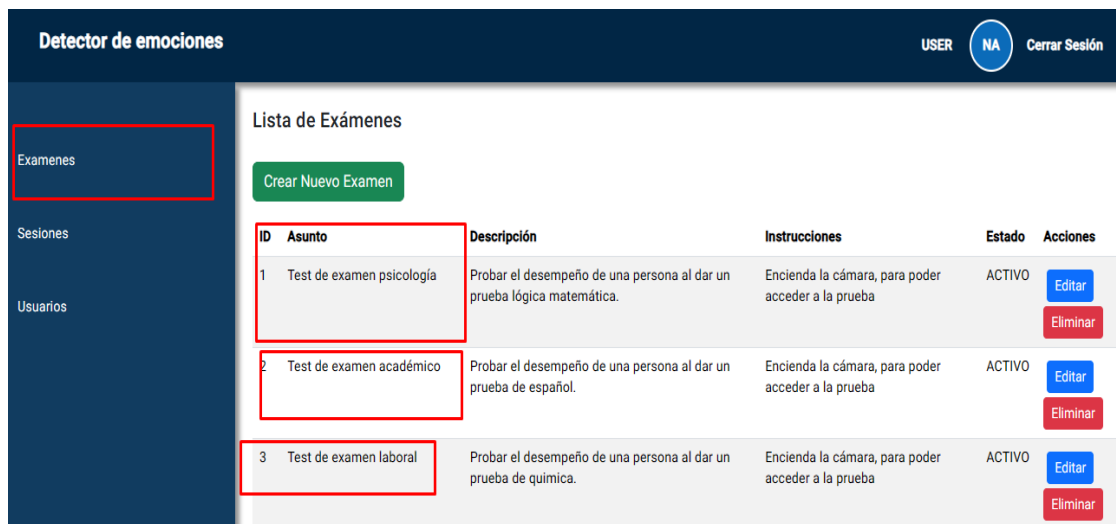


Figura 67. Visualización de la opción exámenes en la interfaz administrativa.

En la pestaña de sesiones, que se presenta en la Figura 68 se visualiza a las personas que han accedido, junto con la información de los test realizados y las observaciones asociadas.



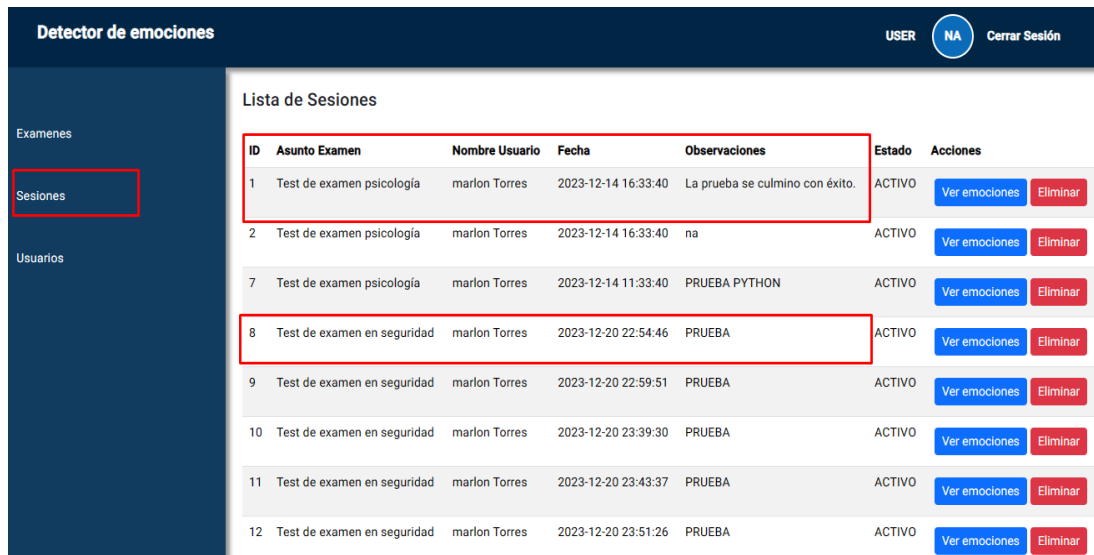


Figura 68. Visualización de la opción sesiones en la interfaz administrativa.

Para realizar un diagnóstico, se implementa la opción de visualizar emociones. Al presionar este botón, se despliega una ventana como se muestra en la Figura 69 que presenta las emociones identificadas en el cliente, junto con gráficos ilustrativos para un análisis más detallado.

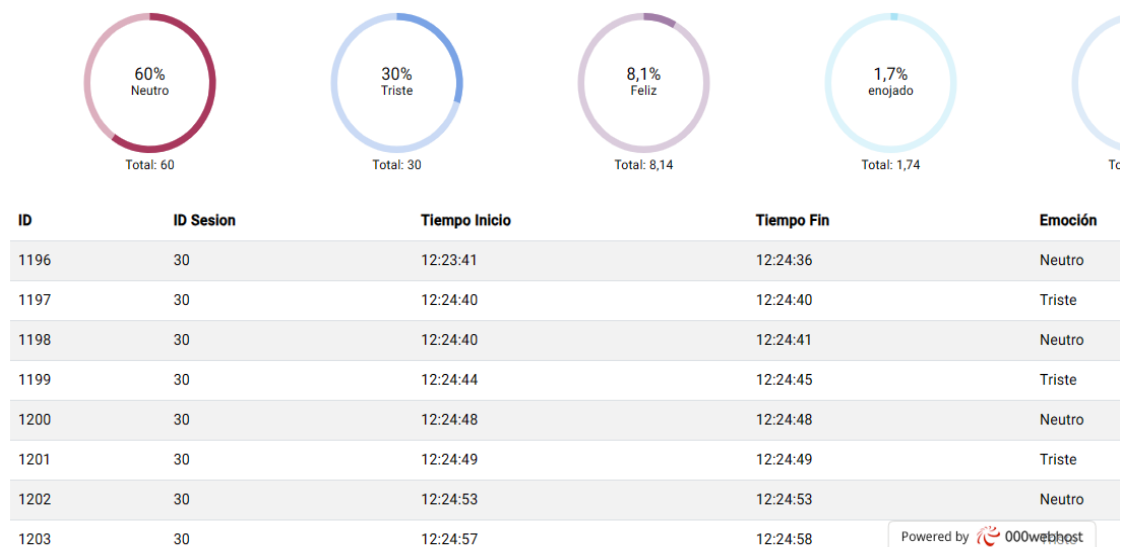


Figura 69. Visualización de los patrones de emociones del usuario.

Finalmente, en la pestaña de usuarios, se visualiza los clientes registrados que tienen acceso a la interfaz de usuario como se visualiza en la Figura 70:

ID	CEDULA	NOMBRE	APELLIDO	ESTADO	CLAVE	TIPO USUARIO	Acciones
1	1804149233	marlon	Torres	a	12345	USUARIO	Editar Eliminar
2	1804149234	Marlon	Maisanche	a	12345	USUARIO	Editar Eliminar
3	1804420816	Daniel	Fiallos	a	12345	ADMINISTRADOR	Editar Eliminar
10	1804420832	Fernando	Fiallos	a	12345	USUARIO	Editar Eliminar
11	1803510864	Isabel	Araujo	a	12345	USUARIO	Editar Eliminar

Figura 70. Lista de usuarios activos.

- **Historial de emociones**

Para obtener un diagnóstico del usuario sometido a estas pruebas, se implementó una validación en la Jetson Nano. A partir de esta validación, los resultados pueden cargarse en el servidor. No obstante, en Jetson, se puede verificar el funcionamiento y la sincronización de la hora para optimizar la eficiencia del envío de datos y reducir errores como se muestra en la Figura 71.

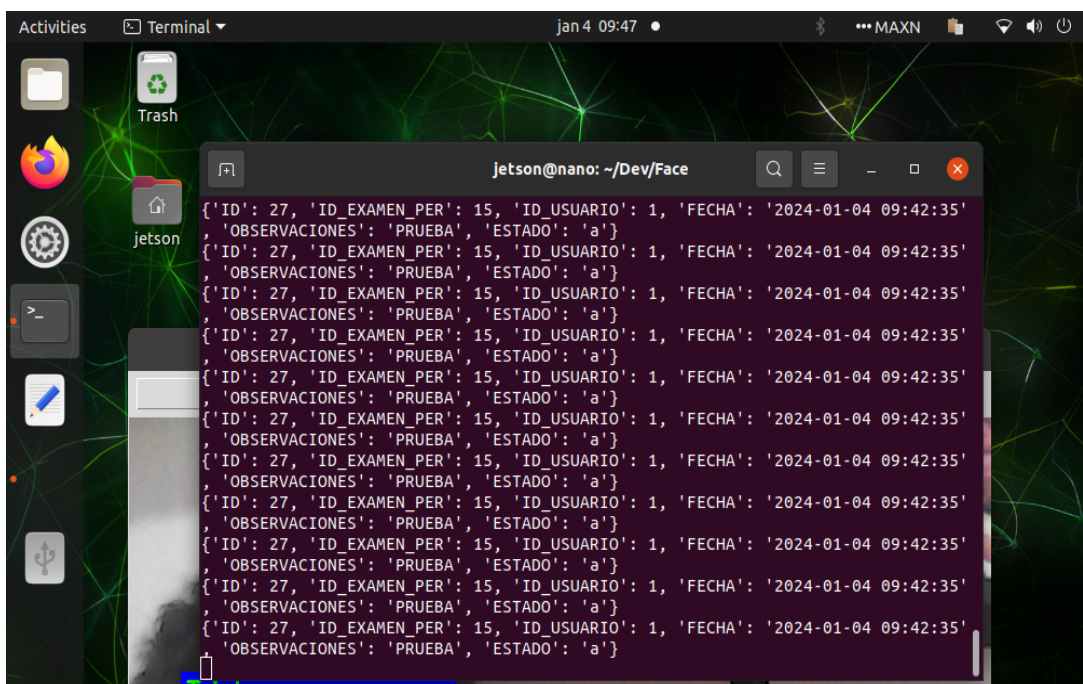
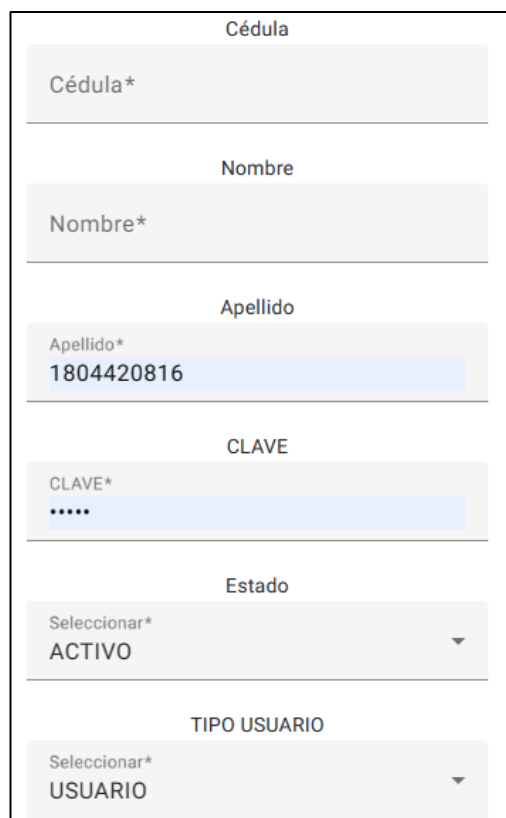


Figura 71. Historial etiquetado de las emociones.

### 3.2.5 Pruebas de funcionamiento

#### Creación de usuario

La interfaz administrativa cuenta con la posibilidad de agregar usuarios y autenticarlos, permitiéndoles acceder a la interfaz de usuario una vez que hayan completado sus datos informativos. Estos datos incluyen el nombre de usuario y la contraseña, tal como se detalla en la Figura 72.



Formulario para crear usuarios. El formulario contiene los siguientes campos:

- Cédula:** Campo de texto con el valor "Cédula\*" y un campo de entrada con el valor "Cédula\*".
- Nombre:** Campo de texto con el valor "Nombre\*" y un campo de entrada con el valor "Nombre\*".
- Apellido:** Campo de texto con el valor "Apellido\*" y un campo de entrada con el valor "1804420816".
- CLAVE:** Campo de texto con el valor "CLAVE\*" y un campo de entrada con el valor "\*\*\*\*\*".
- Estado:** Campo de selección con el valor "Seleccionar\*" y un menú desplegable con el valor "ACTIVO".
- TIPO USUARIO:** Campo de selección con el valor "Seleccionar\*" y un menú desplegable con el valor "USUARIO".

Figura 72. Formulario para crear usuarios.

#### Ingreso en la interfaz

Para acceder a la interfaz del sistema, los usuarios solo necesitan ingresar a través de la interfaz especialmente diseñada. Esta interfaz facilita la interacción del usuario con el sistema de reconocimiento de emociones basado en inteligencia artificial, y ha sido concebida para mejorar la experiencia del usuario al proporcionar una interacción intuitiva. Esto permite realizar de manera eficiente las acciones necesarias para la detección y análisis en tiempo real de los indicadores faciales.

a. Prueba 1

- Activación del proceso

1. **Usuario:** David Palan, autenticado como se observa en la Figura 73.

Crear Nuevo

ID	CEDULA	NOMBRE	APELLIDO	ESTADO	CLAVE	TIPO USUARIO	Acciones
1	1804149233	marlon	Torres	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
2	1804149234	Marlon	Maisanche	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
3	1804420816	Daniel	Fiallos	a	12345	ADMINISTRADOR	<a href="#">Editar</a> <a href="#">Eliminar</a>
10	1804420832	Fernando	Fiallos	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
11	1803510864	Isabel	Araujo	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
12	1804204988	David	Palan	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
13	1850034057	Shirley	Pazmiño	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>

Figura 73. Activación del usuario.

2. **Emociones detectadas:** A continuación, en la Figura 74, Figura 75, Figura 76 y Figura 77 se indican las emociones detectadas por el sistema.

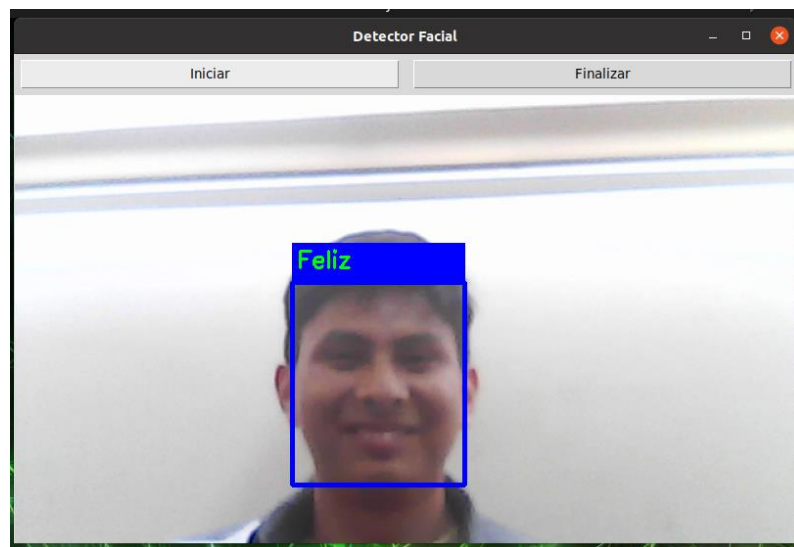


Figura 74. Emoción de felicidad detectada

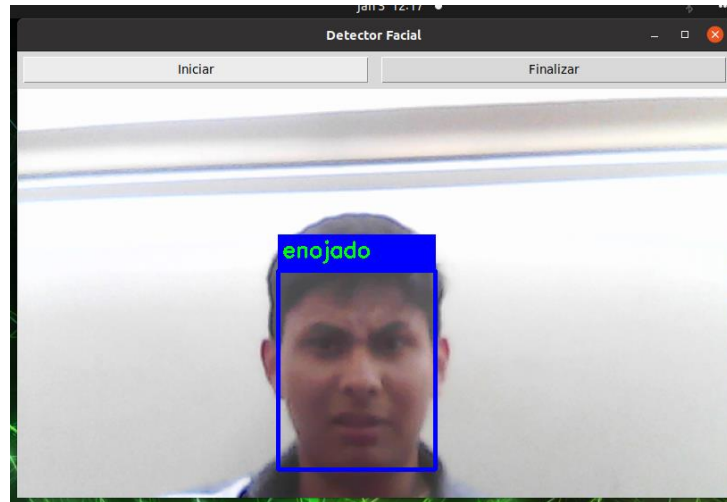


Figura 75. Emoción de enojo detectada.

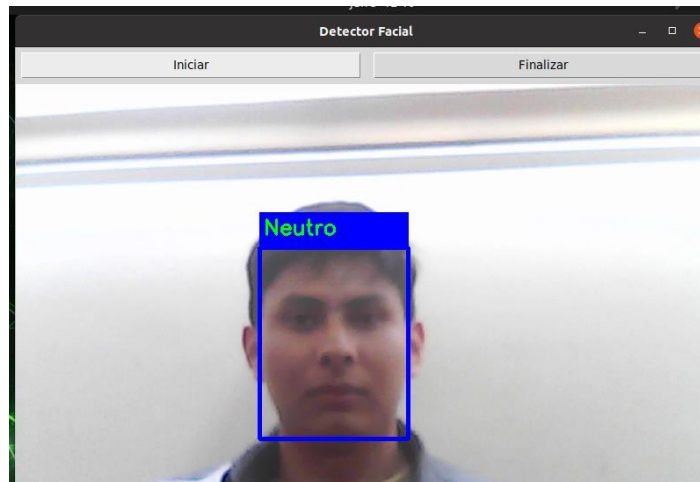


Figura 76. Emoción neutra detectada.

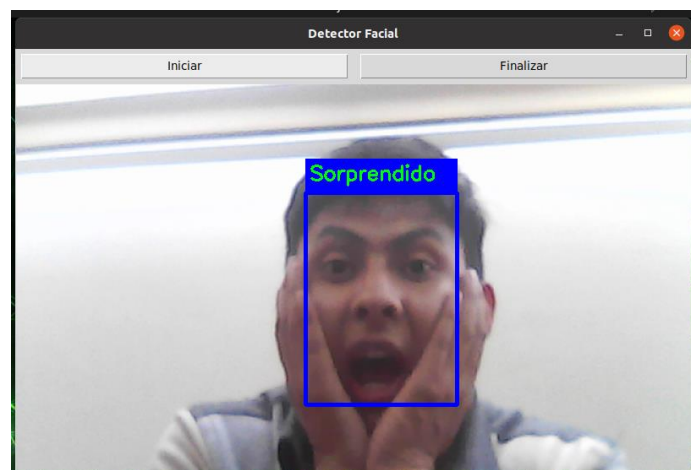


Figura 77. Emoción de sorpresa detectada.

**3. Historial de Emociones.** En la Figura 78 se observa el patrón de emociones del usuario en un lapso.



Figura 78. Patrón de emociones del usuario.

**b. Prueba 2**

- Activación del proceso**

**1. Usuario:** Isabel Araujo, autenticado como se observa en la Figura 79.

Crear Nuevo

ID	CEDULA	NOMBRE	APELLIDO	ESTADO	CLAVE	TIPO USUARIO	Acciones
1	1804149233	marlon	Torres	a	12345	USUARIO	Editar Eliminar
2	1804149234	Marlon	Maisanche	a	12345	USUARIO	Editar Eliminar
3	1804420816	Daniel	Fiallos	a	12345	ADMINISTRADOR	Editar Eliminar
10	1804420832	Fernando	Fiallos	a	12345	USUARIO	Editar Eliminar
11	1803510864	Isabel	Araujo	a	12345	USUARIO	Editar Eliminar
12	1804204988	David	Palan	a	12345	USUARIO	Editar Eliminar
13	1850034057	Shirley	Pazmiño	a	12345	USUARIO	Editar Eliminar

Figura 79. Activación del usuario.

2. **Emociones detectadas.** A continuación, en la Figura 80, Figura 81 y Figura 82 se indican las emociones detectadas por el sistema.

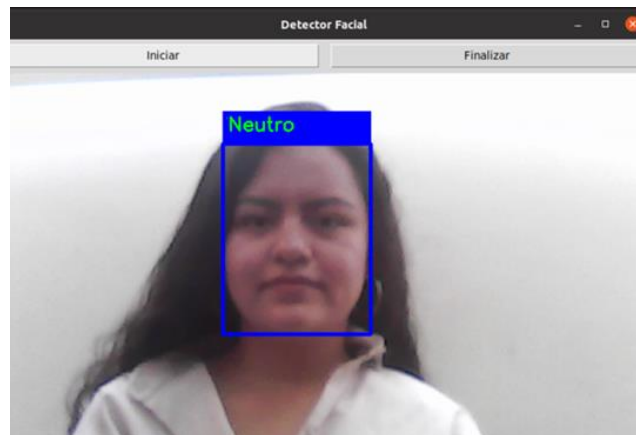


Figura 80. Emoción neutra detectada.

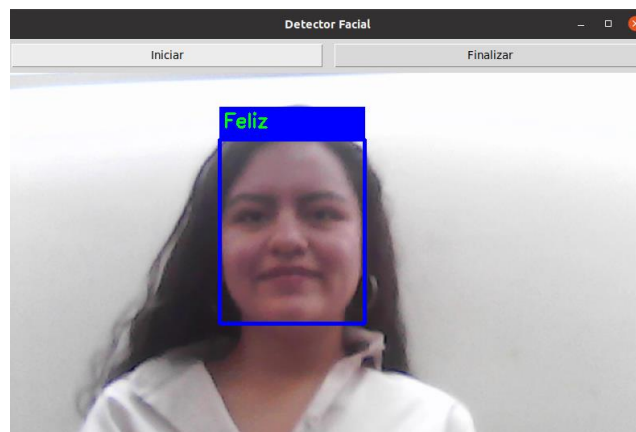


Figura 81. Emoción de felicidad detectada.

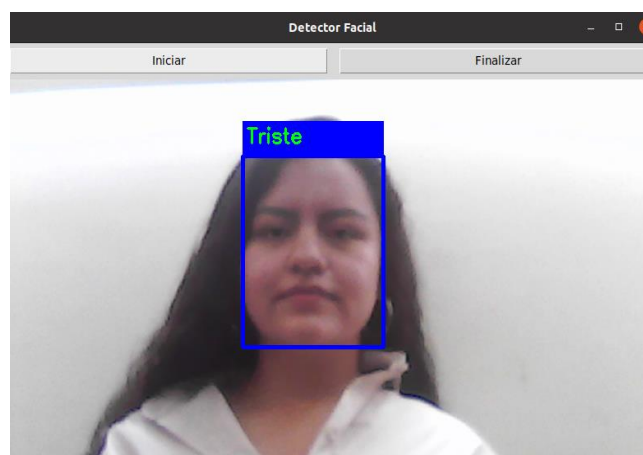


Figura 82. Emoción de tristeza detectada.

3. **Historial de Emociones.** En la Figura 83 se observa el patrón de emociones del usuario en un lapso.

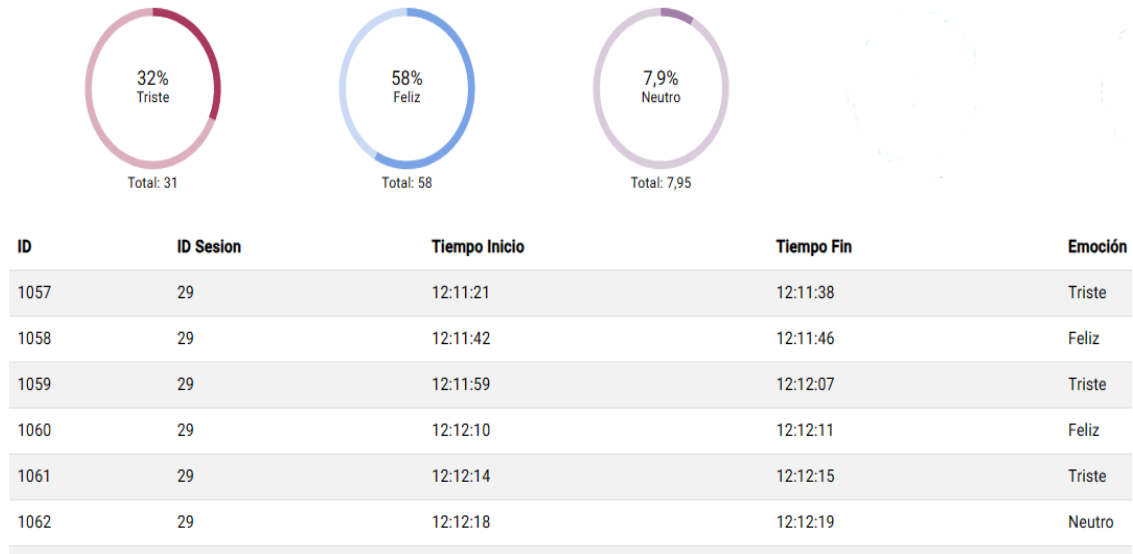


Figura 83. Patrón de emociones del usuario.

### c. Prueba 3

- *Activación del proceso*

1. **Usuario:** Shirley Pazmiño, autenticado como se observa en la Figura 84.

ID	CEDULA	NOMBRE	APELLIDO	ESTADO	CLAVE	TIPO USUARIO	Acciones
1	1804149233	marlon	Torres	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
2	1804149234	Marlon	Maisanche	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
3	1804420816	Daniel	Fiallos	a	12345	ADMINISTRADOR	<a href="#">Editar</a> <a href="#">Eliminar</a>
10	1804420832	Fernando	Fiallos	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
11	1803510864	Isabel	Araujo	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
12	1804204988	David	Palan	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>
13	1850034057	Shirley	Pazmiño	a	12345	USUARIO	<a href="#">Editar</a> <a href="#">Eliminar</a>

Figura 84. Activación del usuario.



**2. Emociones detectadas.** A continuación, en la Figura 85 se indican la emoción detectadas por el sistema.

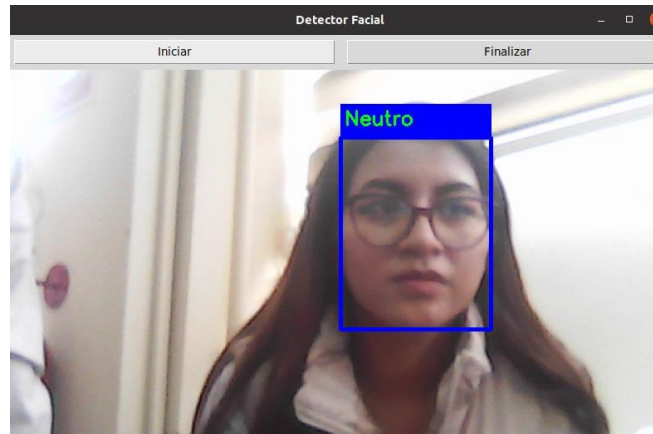


Figura 85. Emoción neutro detectada.

**3. Historial de Emociones.** En la Figura 86 se observa el patrón de emociones del usuario en un lapso.

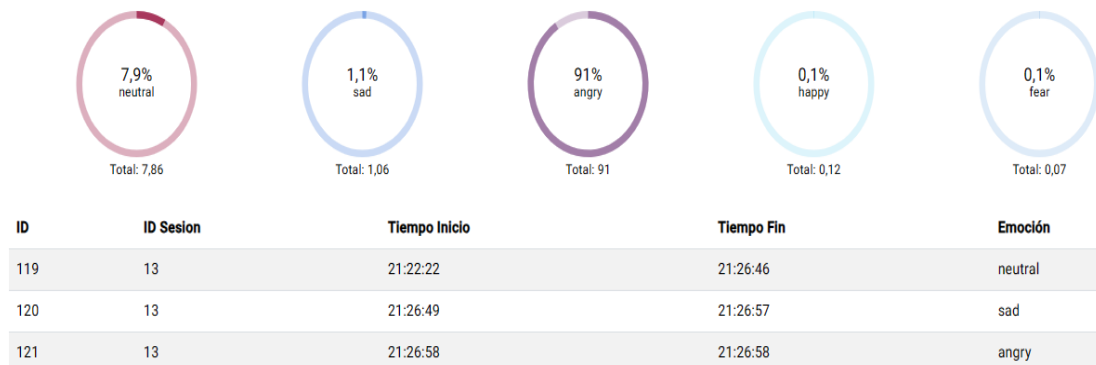


Figura 86. Patrón de emociones del usuario.

### 3.2.6 Resultados

Para interpretar los resultados del sistema de indicadores faciales para la detección de emociones se realizó la matriz de confusión de cada emoción con una muestra de 100 personas.

#### a. *Matriz de confusión de la detección de felicidad.*

La matriz de confusión de la emoción felicidad se realizó mediante el análisis de resultados a 100 personas, que se resume en la Tabla 29.

Tabla 29. Matriz de confusión.

Predicción	Valor positivo	Valor negativo
Predicción positiva	71 (VP)	5 (FN)
Predicción negativa	6 (FP)	18 (VN)

- **Interpretación de resultados.** Los datos fueron recopilados a través de pruebas de funcionamiento utilizando fotografías de individuos, así como con la participación de personas reales. Se considero la emoción que cada persona transmitía durante estas pruebas.

**Verdaderos positivos: 75**

El sistema ha detectado con exactitud a 75 personas que presentaban indicadores faciales de la emoción felicidad.

**Verdaderos negativos: 22**

El sistema ha realizado una identificación precisa de 22 casos de manera acertada que no presentaban signos de felicidad.

**Falsos positivos: 1**

Se registró una instancia en la que el sistema ha categorizado erróneamente a una persona como feliz cuando, de hecho, no lo estaba.

**Falsos negativos: 2**

En dos situaciones, el sistema no ha reconocido correctamente la emoción de felicidad de las personas, indicando que el sistema no detectó una condición que realmente existía.

**b. Matriz de confusión de la detección de tristeza.**

La matriz de confusión de la emoción tristeza se realizó mediante el análisis de resultados a 100 personas, que se resume en la Tabla 30.

Tabla 30. Matriz de confusión.

Predicción	Valor positivo	Valor negativo
Predicción positiva	70 (VP)	3 (FN)
Predicción negativa	2 (FP)	25 (VN)

- **Interpretación de resultados.** Los datos fueron recopilados a través de pruebas de funcionamiento utilizando fotografías de individuos, así como con la participación de personas reales. Se considero la emoción que cada persona transmitía durante estas pruebas.

**Verdaderos positivos: 70**

El sistema ha detectado con exactitud a 70 personas que presentaban indicadores faciales de la emoción tristeza.

**Verdaderos negativos: 25**

El sistema ha identificado 25 casos que no presentaban signos de tristeza.

**Falsos positivos: 2**

Se registró 2 instancias en la que el sistema ha categorizado erróneamente a una persona como triste cuando, de hecho, no lo estaba.

**Falsos negativos: 3**

En 3 situaciones, el sistema no ha reconocido correctamente la emoción de tristeza de las personas, indicando que el sistema no detectó una condición que realmente existía.

**c. Matriz de confusión de la detección neutro**

La matriz de confusión de la emoción neutro se realizó mediante el análisis de resultados a 100 personas, que se resume en la Tabla 31.

Tabla 31. Matriz de confusión.

Predicción	Valor positivo	Valor negativo
Predicción positiva	68 (VP)	7 (FN)
Predicción negativa	4 (FP)	21 (VN)

- **Interpretación de resultados.** Los datos fueron recopilados a través de pruebas de funcionamiento utilizando fotografías de individuos, así como con la participación de personas reales. Se considero la emoción que cada persona transmitía durante estas pruebas.

**Verdaderos positivos: 68**

El sistema ha detectado con exactitud a 68 personas que no presentaban una emoción es decir neutro.

**Verdaderos negativos: 21**

El sistema ha realizado una identificación precisa de 21 casos de manera acertada que no presentaban ninguna emoción.

**Falsos positivos: 4**

Se registró 4 instancias en la que el sistema ha categorizado erróneamente a una persona como neutra cuando, de hecho, no lo estaba.

**Falsos negativos: 7**

En 7 situaciones, el sistema no ha reconocido correctamente la emoción de neutro de las personas, indicando que el sistema no detectó una condición que realmente existía.

**d. Matriz de confusión de la detección de enojo.**

La matriz de confusión de la emoción enojo se realizó mediante el análisis de resultados a 100 personas, que se resume en la Tabla 32.

Tabla 32. Matriz de confusión.

Predicción	Valor positivo	Valor negativo
Predicción positiva	72 (VP)	4 (FN)
Predicción negativa	3 (FP)	21 (VN)

- **Interpretación de resultados.** Los datos fueron recopilados a través de pruebas de funcionamiento utilizando fotografías de individuos, así como con la participación

de personas reales. Se considero la emoción que cada persona transmitía durante estas pruebas.

**Verdaderos positivos: 72**

El sistema ha detectado con exactitud a 72 personas que presentaban indicadores fáciles de la emoción enojo.

**Verdaderos negativos: 21**

El sistema ha realizado una identificación precisa de 21 casos de manera acertada que no presentaban signos de enojo.

**Falsos positivos: 3**

Se registró 3 instancias en la que el sistema ha categorizado erróneamente a una persona como enojada cuando, de hecho, no lo estaba.

**Falsos negativos: 4**

En 4 situaciones, el sistema no ha reconocido correctamente la emoción de ira de las personas, indicando que el sistema no detectó una condición que realmente existía.

*e. Matriz de confusión de la detección de sorpresa*

La matriz de confusión de la emoción sorpresa se realizó mediante el análisis de resultados a 100 personas, que se resume en la Tabla 33.

Tabla 33. Matriz de confusión.

Predicción	Valor positivo	Valor negativo
Predicción positiva	72 (VP)	4 (FN)
Predicción negativa	3 (FP)	21 (VN)

- **Interpretación de resultados.** Los datos fueron recopilados a través de pruebas de funcionamiento utilizando fotografías de individuos, así como con la participación de personas reales. Se considero la emoción que cada persona transmitía durante estas pruebas.

**Verdaderos positivos: 72**

El sistema ha detectado con exactitud a 72 personas que presentaban indicadores fáciles de la emoción sorpresa.

**Verdaderos negativos: 21**

El sistema ha realizado una identificación precisa de 21 casos de manera acertada que no presentaban signos de sorpresa.

**Falsos positivos: 3**

Se registró 3 instancias en la que el sistema ha categorizado erróneamente a una persona como sorprendida cuando, de hecho, no lo estaba.

**Falsos negativos: 4**

En 4 situaciones, el sistema no ha reconocido correctamente la emoción de sorpresa de las personas, indicando que el sistema no detectó una condición que realmente existía.

*f. Matriz de confusión de la detección del miedo*

La matriz de confusión de la emoción miedo se realizó mediante el análisis de resultados a 100 personas, que se resume en la Tabla 34.

Tabla 34. Matriz de confusión.

Predicción	Valor positivo	Valor negativo
Predicción positiva	71 (VP)	4 (FN)
Predicción negativa	6 (FP)	19 (VN)

- **Interpretación de resultados.** Los datos fueron recopilados a través de pruebas de funcionamiento utilizando fotografías de individuos, así como con la participación de personas reales. Se considero la emoción que cada persona transmitía durante estas pruebas.

**Verdaderos positivos: 71**

El sistema ha detectado con exactitud a 71 personas que presentaban indicadores faciales de la emoción miedo.

**Verdaderos negativos: 19**

El sistema ha realizado una identificación precisa de 19 casos de manera acertada que no presentaban signos de miedo.

**Falsos positivos: 6**

Se registró 6 instancias en la que el sistema ha categorizado erróneamente a una persona con miedo cuando, de hecho, no lo estaba.

**Falsos negativos: 4**

En 4 situaciones, el sistema no ha reconocido correctamente la emoción de miedo de las personas, indicando que el sistema no detectó una condición que realmente existía.

**g. Matriz de confusión de la detección de disgusto o asco.**

La matriz de confusión de la emoción disgusto se realizó mediante el análisis de resultados a 100 personas, que se resume en la Tabla 35.

Tabla 35. Matriz de confusión del disgusto.

Predicción	Valor positivo	Valor negativo
Predicción positiva	69 (VP)	5 (FN)
Predicción negativa	6 (FP)	20 (VN)

- **Interpretación de resultados.** Los datos fueron recopilados a través de pruebas de funcionamiento utilizando fotografías de individuos, así como con la participación de personas reales. Se considero la emoción que cada persona transmitía durante estas pruebas.

**Verdaderos positivos: 69**

El sistema ha detectado con exactitud a 69 personas que presentaban indicadores faciales de la emoción disgusto.

**Verdaderos negativos: 20**

El sistema ha realizado una identificación precisa de 20 casos de manera acertada que no presentaban signos de asco.

**Falsos positivos: 6**

Se registró 6 instancias en la que el sistema ha categorizado erróneamente a una persona como disgustada cuando, de hecho, no lo estaba.

**Falsos negativos: 5**

En 5 situaciones, el sistema no ha reconocido correctamente la emoción de disgusto de las personas, indicando que el sistema no detectó una condición que realmente existía.

**3.2.7 Validación de resultados**

*a. Cálculos de la emoción felicidad*

- *Precisión*

$$Precisión = \frac{VP}{(VP + FP)} \quad (1)$$

$$Precisión = \frac{71}{(71 + 6)}$$

$$Precisión = 0,922$$



- **Sensibilidad (Recall)**

$$Recall = \frac{VP}{(VP + FN)} \quad (2)$$

$$Recall = \frac{71}{(71 + 4)}$$

$$Recall = 0,946$$

- **F1-score**

$$f1\_score = 2 * \left( \frac{Precisión * Recall}{Precisión + Recall} \right) \quad (3)$$

$$f1\_score = 2 * \left( \frac{0,922 * 0,946}{0,922 + 0,946} \right)$$

$$f1\_score = 0,934$$

- **Accuracy (exactitud)**

$$Exactitud = \left( \frac{VP + TN}{VP + FP + VN + FN} \right) \quad (4)$$

$$Exactitud = \left( \frac{71 + 19}{71 + 6 + 19 + 4} \right)$$

$$Exactitud = 0,9$$

#### **b. Cálculos de la emoción tristeza**

- **Precisión**

$$Precisión = \frac{VP}{(VP + FP)}$$

$$\text{Precisión} = \frac{70}{(70 + 2)}$$

$$\text{Precisión} = 0,972$$

- **Sensibilidad (Recall)**

$$\text{Recall} = \frac{VP}{(VP + FN)}$$

$$\text{Recall} = \frac{70}{(70 + 3)}$$

$$\text{Recall} = 0,958$$

- **F1-score**

$$f1\_score = 2 * \left( \frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}} \right)$$

$$f1\_score = 2 * \left( \frac{0,972 * 0,958}{0,972 + 0,958} \right)$$

$$f1\_score = 0,964$$

- **Accuracy (exactitud)**

$$\text{Exactitud} = \left( \frac{VP + VN}{VP + FP + VN + FN} \right)$$

$$\text{Exactitud} = \left( \frac{70 + 25}{70 + 2 + 25 + 3} \right)$$

$$\text{Exactitud} = 0,95$$

c. **Cálculos de la emoción neutro.**

- **Precisión**

$$\text{Precisión} = \frac{VP}{(VP + FP)}$$

$$\text{Precisión} = \frac{68}{(68 + 4)}$$

$$\text{Precisión} = 0,944$$

- **Sensibilidad (Recall)**

$$\text{Recall} = \frac{VP}{(VP + FN)}$$

$$\text{Recall} = \frac{68}{(68 + 7)}$$

$$\text{Recall} = 0,906$$

- **F1-score**

$$f1\_score = 2 * \left( \frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}} \right)$$

$$f1\_score = 2 * \left( \frac{0,944 * 0,906}{0,944 + 0,906} \right)$$

$$f1\_score = 0,924$$

- **Accuracy (exactitud)**

$$\text{Exactitud} = \left( \frac{VP + TN}{VP + FP + VN + FN} \right)$$

$$\text{Exactitud} = \left( \frac{68 + 21}{68 + 4 + 21 + 7} \right)$$

$$\text{Exactitud} = 0,89$$

**d. Cálculos de la emoción enojo.**

- **Precisión**

$$\text{Precisión} = \frac{VP}{(VP + FP)}$$

$$\text{Precisión} = \frac{72}{(72 + 3)}$$

$$\text{Precisión} = 0,96$$

- **Sensibilidad (Recall)**

$$\text{Recall} = \frac{VP}{(VP + FN)}$$

$$\text{Recall} = \frac{72}{(72 + 4)}$$

$$\text{Recall} = 0,947$$

- **F1-score**

$$f1\_score = 2 * \left( \frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}} \right)$$

$$f1\_score = 2 * \left( \frac{0,96 * 0,947}{0,96 + 0,947} \right)$$

$$f1\_score = 0,953$$

- **Accuracy (exactitud)**

$$\text{Exactitud} = \left( \frac{VP + VN}{VP + FP + VN + FN} \right)$$

$$\text{Exactitud} = \left( \frac{72 + 21}{72 + 3 + 21 + 4} \right)$$

$$\text{Exactitud} = 0,93$$

e. **Cálculos de la emoción sorpresa**

- **Precisión**

$$\text{Precisión} = \frac{VP}{(VP + FP)}$$

$$\text{Precisión} = \frac{71}{(71 + 6)}$$

$$\text{Precisión} = 0,922$$

- **Sensibilidad (Recall)**

$$\text{Recall} = \frac{VP}{(VP + FN)}$$

$$\text{Recall} = \frac{71}{(71 + 4)}$$

$$\text{Recall} = 0,946$$

- **F1-score**

$$f1\_score = 2 * \left( \frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}} \right)$$

$$f1\_score = 2 * \left( \frac{0,922 * 0,946}{0,922 + 0,946} \right)$$

$$f1\_score = 0,934$$

- **Accuracy (exactitud)**

$$\text{Exactitud} = \left( \frac{VP + TN}{VP + FP + VN + FN} \right)$$

$$\text{Exactitud} = \left( \frac{71 + 19}{71 + 6 + 19 + 4} \right)$$

$$\text{Exactitud} = 0,9$$

**f. Cálculos de la emoción miedo.**

- **Precisión**

$$\text{Precisión} = \frac{VP}{(VP + FP)}$$

$$\text{Precisión} = \frac{71}{(71 + 6)}$$

$$\text{Precisión} = 0,922$$

- **Sensibilidad (Recall).**

$$\text{Recall} = \frac{VP}{(VP + FN)}$$

$$\text{Recall} = \frac{71}{(71 + 4)}$$

$$\text{Recall} = 0,946$$

- **F1-score.**

$$f1\_score = 2 * \left( \frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}} \right)$$

$$f1\_score = 2 * \left( \frac{0,922 * 0,946}{0,922 + 0,946} \right)$$

$$f1\_score = 0,934$$

- **Accuracy (exactitud).**

$$\text{Exactitud} = \left( \frac{VP + TN}{VP + FP + VN + FN} \right)$$

$$\text{Exactitud} = \left( \frac{71 + 19}{71 + 6 + 19 + 4} \right)$$

$$\text{Exactitud} = 0,9$$

**g. Cálculos de la emoción disgusto o asco.**

- **Precisión**

$$\text{Precisión} = \frac{VP}{(VP + FP)}$$

$$\text{Precisión} = \frac{69}{(69 + 6)}$$

$$\text{Precisión} = 0,92$$

- **Sensibilidad (Recall)**

$$\text{Recall} = \frac{VP}{(VP + FN)}$$

$$\text{Recall} = \frac{69}{(69 + 5)}$$

$$\text{Recall} = 0,932$$

- **F1-score**

$$f1\_score = 2 * \left( \frac{\text{Precisión} * \text{Recall}}{\text{Precisión} + \text{Recall}} \right)$$

$$f1\_score = 2 * \left( \frac{0,92 * 0,932}{0,92 + 0,932} \right)$$

$$f1\_score = 0,925$$

- **Accuracy (exactitud)**

$$\text{Exactitud} = \left( \frac{VP + VN}{VP + FP + VN + FN} \right)$$

$$\text{Exactitud} = \left( \frac{69 + 18}{69 + 6 + 20 + 5} \right)$$

$$\text{Exactitud} = 0,90$$

A continuación, se presenta la Tabla 36, que proporciona los valores específicos de los cuatro parámetros establecidos para validar el modelo en relación con cada emoción. Cada fila de la tabla representa una emoción distinta, mientras que las columnas corresponden a los parámetros definidos para medir la eficacia y precisión del modelo en la detección de cada una de estas emociones. La información presentada en esta tabla es fundamental para comprender el desempeño del modelo en diferentes

contextos emocionales, contribuyendo así a una evaluación completa y detallada de su capacidad predictiva.

Tabla 36. Resultados generales de la pruebas del sistema.

Emociones	Precision	Recall	F1-score	Accuracy
Felicidad	0,922	0,946	0,934	0,90
Tristeza	0,972	0,958	0,964	0,95
Neutro	0,944	0,906	0,924	0,89
Enojo	0,96	0,947	0,953	0,93
Sorpresa	0,922	0,946	0,934	0,90
Miedo	0,922	0,946	0,934	0,90
Disgusto	0,92	0,932	0,925	0,90
<b>Total</b>	0,937	0,940	0,938	0,91

Con base en los parámetros de validación analizados, se puede concluir que el sistema de detección cuenta con una elevada precisión y recall, estableciendo su capacidad para determinar las emociones, el parámetro F1-score, indica un balance integral que combina la precision y el recall. La precision del sistema es aceptable, sin embargo, la emoción de disgusto o asco es la que menor valor cuenta, debido a su complejidad para hacer los patrones faciales de dicha emoción de manera natural, por el contrario, las emociones de felicidad, tristeza y neutro son los valores más altos debido a su manera sencilla y natural de expresar estas emociones.

El parámetro de F1-score, ayuda a evaluar de manera porcentual el rendimiento y versatilidad del modelo aplicado en pruebas de entornos reales, este valor se normaliza de 1 a 0 siendo el valor más alto un rendimiento perfecto y el valor más bajo un rendimiento nulo, después de un proceso de calcular el promedio total del sistema del parámetro F1-score dio como resultado un valor de 0,938.

$$\text{Rendimiento del sistema} = F1 - score * 100 \quad (5)$$

$$\text{Rendimiento del sistema} = 0,938 * 100$$

$$\text{Rendimiento del sistema} = \mathbf{93,8\%}$$

Este porcentaje determina el rendimiento del sistema detector de emociones mediante inteligencia artificial y su versatilidad al clasificar las emociones de manera correcta y eficiente, es fundamental comprender que este parámetro será muy variado debido a diversos factores como los entornos reales de prueba y los datos característicos adquiridos.



### 3.2.8 Presupuesto

En el análisis de presupuesto del sistema de detección de emociones mediante el análisis de indicadores faciales empleando inteligencia artificial, se subdivide en dos componentes fundamentales: el presupuesto destinado al diseño y el presupuesto de fabricación. Para obtener el presupuesto de diseño se evaluaron las horas empleadas en el diseño del sistema con el sueldo estipulado por el ministerio de Trabajo de Ecuador de un ingeniero en telecomunicaciones de \$858 dólares. Para determinar el salario por día se considera los 22 días laborables:

$$\text{Salario}_{\text{día}} = \frac{\text{sueldo}}{\text{días}_{\text{laborables}}} \quad (6)$$

$$\text{Salario}_{\text{día}} = \frac{858 \text{ dólares}}{21 \text{ días}}$$

$$\text{Salario}_{\text{día}} = 40,86 \text{ dólares}$$

Para determinar el salario por hora se toma a consideración la jornada laboral por día que son 8 horas en relación con el salario por día:

$$\text{Salario}_{\text{hora}} = \frac{\text{salario}_{\text{día}}}{\text{Horas}_{\text{laborables}}} \quad (7)$$

$$\text{Salario}_{\text{hora}} = \frac{40,86 \text{ dólares}}{8 \text{ horas}}$$

$$\text{Salario}_{\text{hora}} = 5,10 \text{ dólares}$$

Se calcula un lapso de 140 horas destinado al estudio, diseño, ejecución y verificación de funcionamiento, dando como resultado el presupuesto del diseño.

$$\text{Presupuesto total}_{\text{diseño}} = \text{Horas}_{\text{implementación}} \times \text{Salario}_{\text{hora}} \quad (8)$$

$$\text{Presupuesto total}_{\text{diseño}} = 140 \text{ horas} \times 5,10 \text{ dólares}$$

$$Presupuesto\ total_{dise\tilde{n}o} = 714\ d\acute{o}lares$$

A continuaci3n, en la Tabla 37, se proporciona un an3lisis de los costos asociados a los materiales empleados en la implementaci3n del sistema

Tabla 37. Presupuesto de la implementaci3n del proyecto.

Ítem	Recursos	Cantidad	Precio Unitario	Precio Total
1	Kit Jetson Nano	1	\$ 180	\$ 180
2	C3mara web	1	\$ 20	\$ 20
3	Monitor	1	\$ 50	\$ 50
4	Teclado y mouse	1	\$ 12	\$ 12
5	Adaptador HDMI a VGA	1	\$ 10	\$ 10
6	Libros	1	\$ 30	\$ 30
7	Plan web hosting	3	\$ 3,80	\$ 11,40
Subtotal				\$ 313,40
Imprevistos (5%)				\$ 37,56
<b>TOTAL</b>				<b>\$ 350,96</b>

Por 3ltimo, el presupuesto total para la implementaci3n del sistema se determina de la siguiente manera:

$$Presupuesto\ total_{sist} = Presupuesto_{dis} + Presupuesto_{imple} \quad (9)$$

$$Presupuesto\ total_{sistema} = 714\ d\acute{o}lares + 350,96\ d\acute{o}lares$$

$$Presupuesto\ total_{sistema} = 1064,96\ d\acute{o}lares$$

## CAPÍTULO IV. CONCLUSIONES Y RECOMENDACIONES

### a. Conclusiones

- La investigación sobre las emociones y su influencia en diversos contextos sociales ha sido esencial para la identificación de indicadores faciales y patrones que nos indican la complejidad del rostro. Esta caracterización detallada ha proporcionado la base conceptual necesaria para la búsqueda de un conjunto de datos integral y enriquecido. La diversidad de este conjunto, compuesto por imágenes en escala de grises con rostros rotados y distorsionados en diferentes ángulos, ha desempeñado un papel fundamental en el proceso de entrenamiento del algoritmo de inteligencia artificial impulsándolo a aprender patrones emocionales de manera más profunda y generalizada. Este enfoque técnico ha elevado la capacidad del modelo para reconocer y clasificar emociones en situaciones diversas.
- La evaluación de algoritmos de inteligencia artificial que permiten ser altamente eficaces en tareas específicas relacionadas con el procesamiento de imágenes ha contribuido a seleccionar Redes neuronales convolucionales (CNN) para el diseño del algoritmo. La arquitectura de las CNN y el uso de bibliotecas como OpenCV, Tensor Flow Y Keras son complementos fundamentales para el desarrollo de un sistema preciso y eficiente en la identificación de rostros y la determinación de emociones. Para el procesamiento de cada frames se utiliza la microcomputadora Jetson nano la cual ayuda adquirir frames en tiempo real mediante el uso de una cámara demostrando un rendimiento sólido y eficiente para el procesamiento en tiempo real, con respuestas de 2 milisegundos en video.
- La implementación de una interfaz intuitiva para los usuarios ha demostrado ser precisa y fácil de observar la detección de la emoción, además la interfaz administrativa ha comprobado ser una herramienta de gestión de datos versátil. Además de cumplir con el objetivo de proporcionar un seguimiento eficiente del historial de emociones, esta interfaz ha elevado la toma de decisiones a un

nivel superior. Los dashboards informativos presentan de manera clara y concisa los patrones emocionales a lo largo del tiempo, brindando información valiosa para la comprensión profunda de las reacciones emocionales y posibilitando ajustes estratégicos en el sistema.

- Los resultados reales mediante la utilización de un conjunto de 100 personas de prueba derivan a una matriz de confusión, que mediante su análisis respaldan de manera contundente la eficacia del sistema, con datos muy eficiente en la tasa de precisión del 98.6%, la sensibilidad del 96%, y la exactitud general del 96% son indicativos de la capacidad del sistema para identificar de manera precisa los indicadores faciales para una correcta clasificación de emociones en tiempo real. Estos hallazgos confirman que el sistema de detección de emociones mediante el análisis de indicadores faciales empleando inteligencia artificial cumple con los objetivos definidos y genera la certeza requerida para su despliegue en entornos prácticos del mundo real.

#### **b. Recomendaciones**

- Es aconseja priorizar el reconocimiento preciso de características o patrones faciales como primer paso en el proceso, ya que estas características forman la base para la identificación y clasificación de emociones. La calidad de la base de datos para el entrenamiento influye directamente en la eficacia del sistema. Dicha base de datos debe incluir una diversidad representativa de rostros, expresiones y contextos. La variedad en la base de datos garantizará una capacidad de generalización del modelo en una amplia gama de situaciones del mundo real.
- Se recomienda realizar ajustes continuos en el modelo de Redes Neuronales Convolucionales (CNN) para evitar el sobreajuste. La optimización debería centrarse en la eficiencia del modelo, minimizando la redundancia de parámetros innecesarios y asegurando una capacidad de entrenamiento eficiente.

- Al instalar dependencias en dispositivos como la Jetson Nano, se debe tener precaución debido a las limitaciones del procesador. Se sugiere utilizar modelos optimizados y evitar la inclusión de modelos innecesariamente pesados. Se deberá instalar una arquitectura adecuada y que sea compatible con las bibliotecas de TensorFlow y Keras que no ocupen mucho espacio, una tarjeta optimizada contribuirá a la eficiencia del procesamiento, mejorando el rendimiento general del sistema sobre todo al momento de recolectar los frames del video en tiempo real.
- En la creación de APIs para la transferencia de datos al servidor, es recomendable incluir información de fecha y hora junto con los datos de reconocimiento emocional. Esto asegura una trazabilidad temporal precisa y evita problemas de desincronización que podrían surgir en el sistema, garantizando la consistencia en la recopilación de datos en tiempo real.
- Se sugiere llevar a cabo pruebas suplementarias y evaluaciones detalladas con el propósito de obtener resultados más precisos y representativos sobre la versatilidad y del rendimiento global del sistema en una variedad de entornos y condiciones operativas. Esta iniciativa busca validar la respuesta, la robustez y la capacidad de adaptación del sistema en diversas situaciones, asegurando así su confiabilidad y efectividad en escenarios de aplicación del mundo real.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] L. Florido, «El estudio socio-histórico de las emociones y los sentimientos en las Ciencias Sociales del siglo XX», *Open Editions Journals*, ene. 2019, [En línea]. Disponible en: <https://journals.openedition.org/revestudsoc/9762>
- [2] M. García, «Detección online de trastornos emocionales en población clínica de niños y adolescentes», vol. 4, n.º 3, pp. 35-45, 3 de diciembre de 2019.
- [3] J. Rodríguez, *Epidemiología de los trastornos mentales en América Latina y el Caribe*. Washington, D.C.: OPS, 2020. [En línea]. Disponible en: <https://iris.paho.org/bitstream/handle/10665.2/740/9789275316320.pdf>
- [4] E. Ortega, «Epidemiología de la morbilidad psiquiátrica en Ecuador», vol. 22, n.º 3, pp. 50-62, 20 de agosto de 2021.
- [5] Mónica Andrago y Pablo Recalde, «Uso de reconocimiento facial de emociones basado en técnicas de Deep Learning para el mejoramiento de la educación.», *UNIVERSIDAD TECNOLÓGICA ISRAEL*, 15 de marzo de 2019. [En línea]. Disponible en: <http://repositorio.uisrael.edu.ec/bitstream/47000/2297/1/UISRAEL-EC-MASTER%20-%20TELEM-378.242-2019-013.pdf>
- [6] María Medina y Gonzalo Amaya, «INTELIGENCIA EMOCIONAL ARTIFICIAL», Universidad Jorge Tadeo Lozano, Bogotá, 2020. [En línea]. Disponible en: <https://expeditiorepositorio.utadeo.edu.co/bitstream/handle/20.500.12010/16617/INTELIGENCIA%20EMOCIONAL%20ARTIFICIAL%20final.docx.pdf?sequence=1&isAllowed=y>
- [7] J. Pinzón y Nydia Sandoval, «DESARROLLO DE SISTEMA PARA LA DETECCIÓN DE EMOCIONES UTILIZANDO VISIÓN ARTIFICIAL E INTELIGENCIA ARTIFICIAL APLICADO A ENTORNOS EDUCATIVOS.», Universidad de Pamplona, España, 2020. [En línea]. Disponible en:

[http://repositoriodspace.unipamplona.edu.co/jspui/bitstream/20.500.12744/688/1/Pinzon\\_2020\\_TG.pdf](http://repositoriodspace.unipamplona.edu.co/jspui/bitstream/20.500.12744/688/1/Pinzon_2020_TG.pdf)

[8] M. Briones y Marco Moreno, «Detección de emociones usando inteligencia artificial para la recomendación personalizada de bebidas», INSTITUTO POLITECNICO NACIONAL ´, México, 2022. [En línea]. Disponible en: <https://www.saber.cic.ipn.mx/SABERv3/Repositorios/webVerArchivo/51635/1>

[9] E. Flores y J. Livia, «Análisis de sentimientos con inteligencia artificial para mejorar el proceso enseñanza-aprendizaje en el aula virtual», *Dialnet*, vol. 53, n.º 2, p. 16, ago. 2023.

[10] L. Garrido, «Emociones ¿Qué Son Y Para Qué Sirven?», *Psicológicamente*. [En línea]. Disponible en: <https://www.psicologicamenteonline.com/emociones-que-son-y-para-que-sirven/>

[11] V. Idígoras y A. Trefois, «Emociones básicas: Qué y cuáles son las emociones primarias y secundarias», *El Perruco*. [En línea]. Disponible en: <https://www.elperruco.com/emociones-basicas-que-cuales-son-emociones-primarias-secundarias/>

[12] E. Azanza, «¿Qué es el reconocimiento facial?», *Veridas*. [En línea]. Disponible en: <https://veridas.com/es/reconocimiento-facial/#:~:text=Los%20sistemas%20de%20reconocimiento%20facial%20basados%20en%20Inteligencia%20Artificial%20permiten,una%20precisión%20mayor%20al%2099%25.>

[13] I. González, «Qué son los análisis de expresión facial y cómo funcionan», *Neuromarketing*. [En línea]. Disponible en: <https://neuromarketing.la/2016/12/los-analisis-expresion-facial-funcionan/>

[14] J. Mateos, «Los músculos de la cara y sus funciones», *Jennifer Mateos Logopedia*. [En línea]. Disponible en: <https://jennifermateoslogopedia.es/los-musculos-de-la-cara-y-sus->

- [15] O. Sánchez, «Procesamiento Digital de Señales», Monografías. [En línea]. Disponible en: <https://www.monografias.com/trabajos95/procesamiento-digital-de-senales/procesamiento-digital-de-senales>
- [16] E. Bertran, *Procesado digital de señales*, vol. 1. Universidad Politécnica de Cataluña. [En línea]. Disponible en: <https://upcommons.upc.edu/bitstream/handle/2099.3/36546/9788498802597.pdf?sequence=5>
- [17] E. Zabala, «Sistema Digital de Adquisición de Datos». 27 de septiembre de 2016. [En línea]. Disponible en: <https://es.wikipedia.org/wiki/Archivo:DigitalDAQv3ESPANOL.pdf>
- [18] P. Pérez y M. Valente, «Fundamentos básicos del procesamiento de imágenes», Famaf. [En línea]. Disponible en: <https://www.famaf.unc.edu.ar/~pperez1/manuales/cim/cap2.html>
- [19] A. Rodríguez, «El álgebra lineal y el procesamiento digital de imágenes», Nibcode. [En línea]. Disponible en: <https://www.nibcode.com/es/blog/1135/algebra-lineal-y-el-procesamiento-digital-de-imagenes-parte-I>
- [20] Á. Martínez, «Bases Teóricas, Digitalización y Análisis de Imágenes». 18 de mayo de 2019. [En línea]. Disponible en: <https://web.archive.org/web/20100619121419/http://www.seap.es/telepatologia/telepatologia03.pdf>
- [21] B. Escalante, «Procesamiento Digital de Imágenes». 18 de septiembre de 2006. [En línea]. Disponible en: [http://lapi.fi-p.unam.mx/wp-content/uploads/PDI\\_Cap1\\_Introduccion.pdf](http://lapi.fi-p.unam.mx/wp-content/uploads/PDI_Cap1_Introduccion.pdf)
- [22] C. Cintas y C. Delrieux, «Procesamiento Inteligente de Imágenes, posicionamiento de vértebras mediante Landmarks y Redes Neuronales». 04 de 2019. [En línea]. Disponible en: [http://sedici.unlp.edu.ar/bitstream/handle/10915/18846/Documento\\_completo.pdf%3Fsequence%3D1](http://sedici.unlp.edu.ar/bitstream/handle/10915/18846/Documento_completo.pdf%3Fsequence%3D1)



- [23] J. Ponce y A. Torres, *Inteligencia Artificial*, 1.<sup>a</sup> ed. Proyecto Latín, 2018. [En línea]. Disponible en: [https://www.researchgate.net/publication/269466259\\_Inteligencia\\_Artificial](https://www.researchgate.net/publication/269466259_Inteligencia_Artificial)
- [24] B. Marr, «¿Qué es la visión artificial?», IBM. [En línea]. Disponible en: <https://www.ibm.com/es-es/topics/computer-vision>
- [25] M. Reynoso, «Reconocimiento facial», Lisa Institute. [En línea]. Disponible en: <https://www.lisainstitute.com/blogs/blog/reconocimiento-facial-como-funciona-quien-utiliza>
- [26] J. Domínguez, «¿Qué es el reconocimiento facial? Usos y evolución», Mobbeel. [En línea]. Disponible en: <https://www.mobbeel.com/reconocimiento-facial/>
- [27] S. Domínguez, «Reconocimiento facial mediante el Análisis de Componentes Principales (PCA)», Escuela Técnica Superior de Ingeniería Universidad de Sevilla, España, 2017.
- [28] C. Murray, «Creación de una canalización de reconocimiento facial con aprendizaje profundo en Tensorflow», HackerNoon. [En línea]. Disponible en: <https://hackernoon.com/es/construccion-de-una-canalizacion-de-reconocimiento-facial-con-aprendizaje-profundo-en-tensorflow-66e7645015b8>
- [29] F. Calisaya, «Patrones faciales», Universidad privada de Tacna, Perú, 2019. [En línea]. Disponible en: <https://repositorio.upt.edu.pe/handle/20.500.12969/101>
- [30] A. Méndez, «Introducción a Machine Learning», *Cinvestav Unidad Guadalajara*, jul. 2020, [En línea]. Disponible en: <https://unidad.gdl.cinvestav.mx/doc/investigacion/computacion/Introduccion-Machine-Learning.pdf>
- [31] C. Vargas, «Aplicaciones de Machine Learning», Tryccore. [En línea]. Disponible en: <https://trycore.co/transformacion-digital/aplicaciones-machine-learning-en-organizaciones/>

- [32] L. Acosta, «Tipos de aprendizaje en Machine Learning». [En línea]. Disponible en: <https://telefonicatech.com/blog/que-algoritmo-elegir-en-ml-aprendizaje>
- [33] J. Ortega, *Big data, machine learning y data science in Python*. Bogotá: Ediciones de la U, 2023.
- [34] C. Ruíz, «Redes Neuronales: Conceptos Básicos y Aplicaciones.» Universidad Tecnológica Nacional, 8 de marzo de 2021. [En línea]. Disponible en: [https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5\\_anio/orientadora1/monografias/matich-redesneuronales.pdf](https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monografias/matich-redesneuronales.pdf)
- [35] E. Mohamed, «Facial Expression Classification», Kaggle. [En línea]. Disponible en: <https://www.kaggle.com/code/essammohamed4320/facial-expression-classification/notebook>
- [36] C. Ortega, «Modelos de inteligencia artificial: Qué es, tipos y el papel de los datos sintéticos». [En línea]. Disponible en: <https://www.questionpro.com/blog/es/modelos-de-inteligencia-artificial/>
- [37] M. Arrimada, «Emociones neutras: qué son, para qué sirven y cómo nos afectan», *Psicología y Mente*. [En línea]. Disponible en: <https://psicologiymente.com/psicologia/emociones-neutras>
- [38] H. Mallitasig, «Paul Ekman y las microexpresiones faciales de las emociones.» Pontificia Universidad Católica del Ecuador, 08/092018. [En línea]. Disponible en: <https://repositorio.pucesa.edu.ec/bitstream/123456789/2691/1/Microexpresiones%20faciales.pdf>
- [39] A. Paredes, *Qué es una placa SBC 1.1*. Scibd, 2021. [En línea]. Disponible en: <https://es.scribd.com/document/595182324/Que-es-una-placa-SBC-1-1>
- [40] O. González, «Todo lo que necesitas saber sobre Jetson Nano de NVIDIA», *BricoGeek*. [En línea]. Disponible en: <https://blog.bricogeek.com/noticias/robotica/todo-lo-que-necesitas-saber-sobre-jetson-nano-de->



- [48] M. Coppola, «Qué es un hosting: características, tipos y ejemplos», HubSpot. [En línea]. Disponible en: <https://blog.hubspot.es/website/hosting>
- [49] D. Vera, «000Webhost la mejor manera de tener un alojamiento web», NIIXER. [En línea]. Disponible en: <https://niixer.com/index.php/2021/04/14/000webhost-la-mejor-manera-de-tener-un-alojamiento-web/>
- [50] R. Weisheim, «25 herramientas para programar con inteligencia artificial: cómo la IA puede optimizar el desarrollo web», Hostinger. [En línea]. Disponible en: [https://www.hostinger.es/tutoriales/programar-con-inteligencia-artificial?ppc\\_campaign=google\\_search\\_generic\\_hosting\\_all&bidkw=defaultkeyword&lo=1005380&gad\\_source=1&gclid=Cj0KCQiAhc-sBhCEARIsAOVwHuSFO2wfilYP5DSxKDNMP1cU1-gk6S8H9PCNWqzXHokTyTgw3EwSwfcaAnhUEALw\\_wcB](https://www.hostinger.es/tutoriales/programar-con-inteligencia-artificial?ppc_campaign=google_search_generic_hosting_all&bidkw=defaultkeyword&lo=1005380&gad_source=1&gclid=Cj0KCQiAhc-sBhCEARIsAOVwHuSFO2wfilYP5DSxKDNMP1cU1-gk6S8H9PCNWqzXHokTyTgw3EwSwfcaAnhUEALw_wcB)
- [51] M. Tafura, «6 herramientas prácticas para experimentar IA», aiUTEC. [En línea]. Disponible en: <https://aichallenge.utec.edu.uy/community/ai/6-herramientas-practicas-para-experimentar-la-ia/#:~:text=Tensorflow%20playground%20es%20otra%20herramienta,tareas%20de%20clasificación%20de%20datos.>
- [52] «JupyterLab Información general de IU», Experience Platform. [En línea]. Disponible en: <https://experienceleague.adobe.com/docs/experience-platform/data-science-workspace/jupyterlab/overview.html?lang=es#:~:text=JupyterLab%20es%20una%20interfaz%20de,Jupyter%20Notebooks%2C%20código%20y%20datos.>
- [53] «¿Qué es Amazon SageMaker?», AWS. [En línea]. Disponible en: [https://docs.aws.amazon.com/es\\_es/sagemaker/latest/dg/whatis.html](https://docs.aws.amazon.com/es_es/sagemaker/latest/dg/whatis.html)
- [54] T. Santos, «Google Colab: ¿qué es y cómo usarlo?», Alura Cursos. [En línea]. Disponible en: <https://www.aluracursos.com/blog/google-colab-que-es-y-como-usarlo>

[55] «DATA SHEET NVIDIA Jetson Nano System-on-Module». NVIDIA. [En línea]. Disponible en: [https://developer.download.nvidia.com/assets/embedded/secure/jetson/Nano/docs/JetsonNano\\_DataSheet\\_DS09366001v1.1.pdf?mPUJA9mXi8EcpECMLQ9Q7dhGCqPbkuZpCTEdH-lZ-UIaAasxvII4Pcr7gybjjua\\_FwPOZkW5q-DiZYxu-ND2plSHOtSNzp-Cse-hPE6kwh7fyn58bFbtYlxOjsMRjLmfC05TX54dFXfOhpvQT85ozt-aNPuETKAmzJfFjUj0XRP0brjVm\\_Ns0bFEVzBcA==&t=eyJscjI6ImdzZW8iLCJsc2QiOiJodHRwczovL3d3dy5nb29nbGUuY29tLyJ9](https://developer.download.nvidia.com/assets/embedded/secure/jetson/Nano/docs/JetsonNano_DataSheet_DS09366001v1.1.pdf?mPUJA9mXi8EcpECMLQ9Q7dhGCqPbkuZpCTEdH-lZ-UIaAasxvII4Pcr7gybjjua_FwPOZkW5q-DiZYxu-ND2plSHOtSNzp-Cse-hPE6kwh7fyn58bFbtYlxOjsMRjLmfC05TX54dFXfOhpvQT85ozt-aNPuETKAmzJfFjUj0XRP0brjVm_Ns0bFEVzBcA==&t=eyJscjI6ImdzZW8iLCJsc2QiOiJodHRwczovL3d3dy5nb29nbGUuY29tLyJ9)

## ANEXOS

### Anexo A. Categorización de variables

En la Figura 87 y Figura 88 se presenta las diferentes variables determinadas para el desarrollo de este proyecto de investigación



Figura 87. Categorización variable 1.

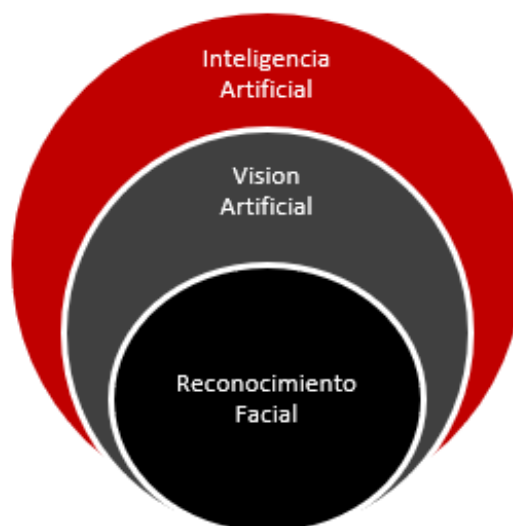


Figura 88. Categorización variable 2

En la Figura 89 y Figura 90 se presentan las constelaciones de ideas de las variables consideradas para el proyecto.

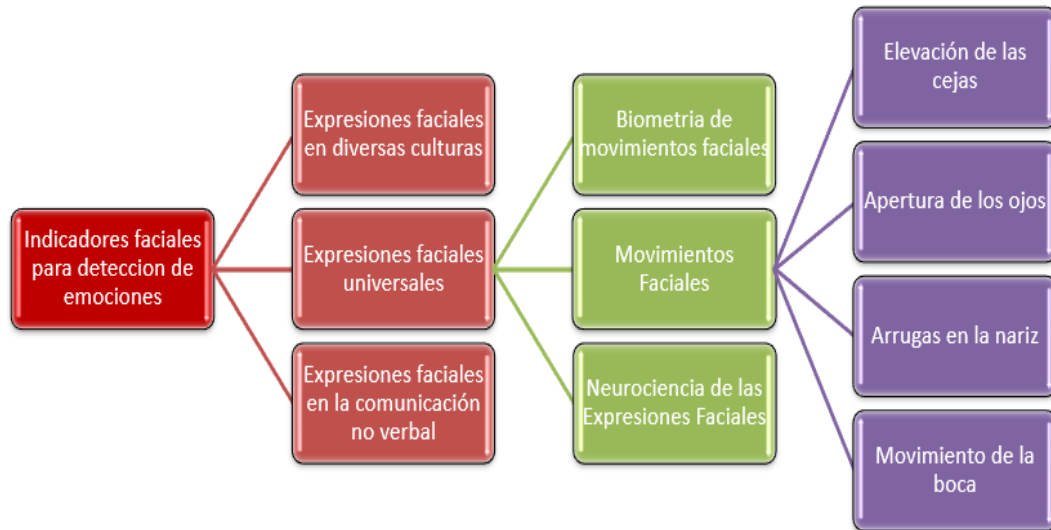


Figura 89. Constelación de ideas variable 1.

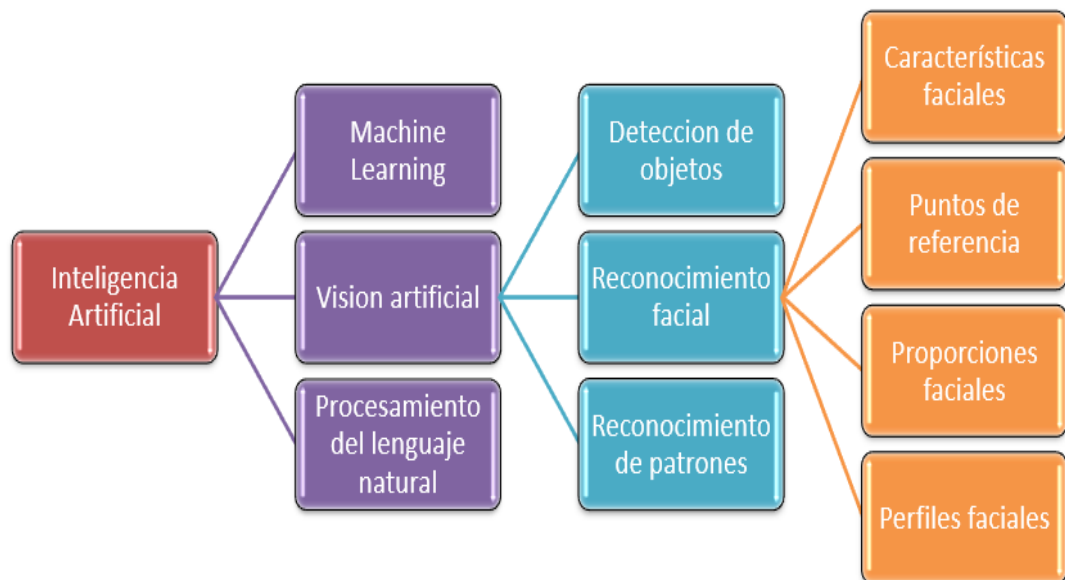



Figura 90. Constelación de ideas variable 2.

## Anexo B. NVIDIA JETSON NANO MODULE

En la Figura 91 y Figura 92 se describen las especificaciones técnicas de la microcomputadora Jetson Nano.



**DATA SHEET**

### NVIDIA Jetson Nano System-on-Module

Maxwell GPU + ARM Cortex-A57 + 4GB LPDDR4 + 16GB eMMC

**Maxwell GPU**<sup>®</sup>  
128-core GPU | End-to-end lossless compression | Tile Caching | OpenGL<sup>®</sup> 4.6 | OpenGL ES 3.2 | Vulkan™ 1.1 | CUDA<sup>®</sup> | OpenGL ES Shader Performance (up to): 512 GFLOPS (FP16)  
Maximum Operating Frequency: 921MHz

**CPU**  
ARM<sup>®</sup> Cortex<sup>®</sup>-A57 MPCore (Quad-Core) Processor with NEON Technology | L1 Cache: 48KB L1 instruction cache (I-cache) per core; 32KB L1 data cache (D-cache) per core | L2 Unified Cache: 2MB | Maximum Operating Frequency: 1.43GHz

**Audio**  
Industry standard High Definition Audio (HDA) controller provides a multichannel audio path to the HDMI interface.

**Memory**  
Dual Channel | System MMU | Memory Type: 4ch x 16-bit LPDDR4 | Maximum Memory Bus Frequency: 1600MHz | Peak Bandwidth: 25.6 GB/s | Memory Capacity: 4GB

**Storage**  
eMMC 5.1 Flash Storage | Bus Width: 8-bit | Maximum Bus Frequency: 200MHz (HS400) | Storage Capacity: 16GB

**Boot Sources**  
eMMC and USB (recovery mode)

**Networking**  
10/100/1000 BASE-T Ethernet | Media Access Controller (MAC)

**Display Controller**  
Two independent display controllers support DSI, HDMI, DP, eDP:  
MIPI-DSI (1.5Gbps/lane): Single x2 lane | Maximum Resolution: 1920x960 at 60Hz (up to 24bpp)  
HDMI 2.0a/b (up to 6Gbps) | DP 1.2a (HBR2 5.4 Gbps) | eDP 1.4 (HBR2 5.4Gbps) | Maximum Resolution (DP/eDP/HDMI): 3840 x 2160 at 60Hz (up to 24bpp)

**Clocks**  
System clock: 38.4MHz | Sleep clock: 32.768kHz | Dynamic clock scaling and clock source selection

**Multi-Stream HD Video and JPEG**

**Video Decode**

- H.265 (Main, Main 10): 2160p 60fps | 1080p 240fps
- H.264 (BP/MP/HP/Stereo SEI half-res): 2160p 60fps | 1080p 240fps
- H.264 (MVC Stereo per view): 2160p 30fps | 1080p 120fps
- VP9 (Profile 0, 8-bit): 2160p 60fps | 1080p 240fps
- VP8: 2160p 60fps | 1080p 240fps
- VC-1 (Simple, Main, Advanced): 1080p 120fps | 1080i 240fps
- MPEG-2 (Main): 2160p 60fps | 1080p 240fps | 1080i 240fps

**Video Encode**

- H.265: 2160p 30fps | 1080p 120fps
- H.264 (BP/MP/HP): 2160p 30fps | 1080p 120fps
- H.264 (MVC Stereo per view): 1440p 30fps | 1080p 60fps
- VP8: 2160p 30fps | 1080p 120fps
- JPEG (Decode and Encode): 600 MP/s

Figura 91. Data sheet de la Jetson Nano [55].



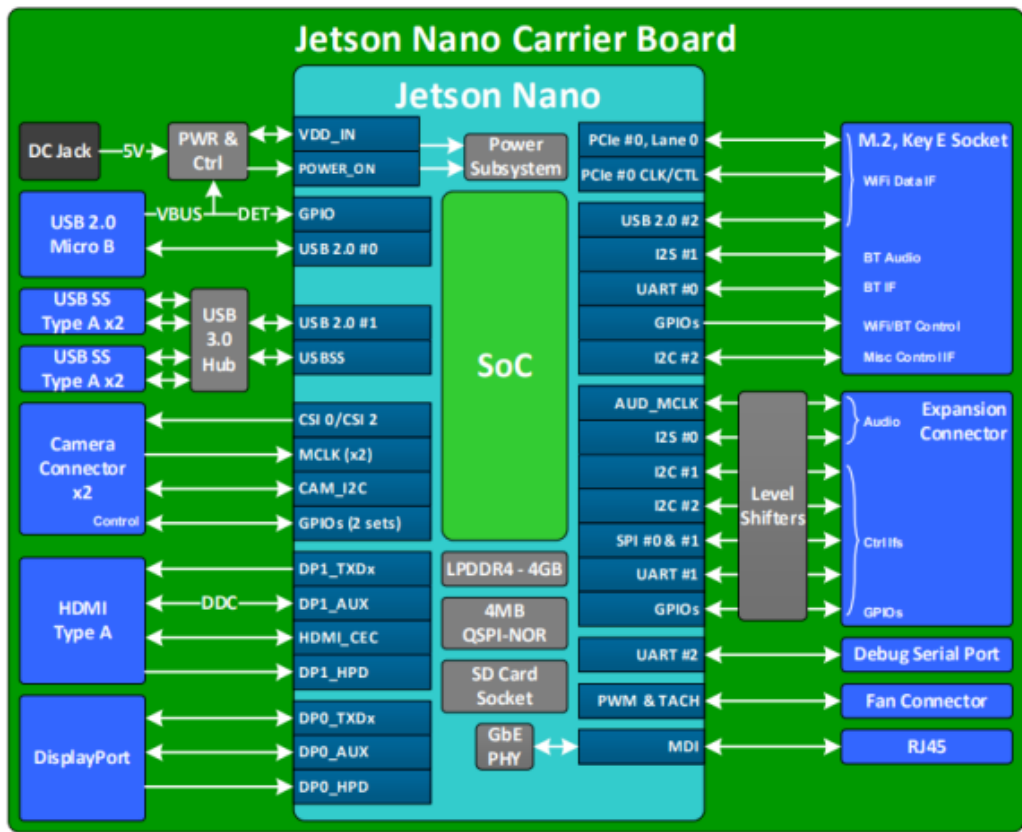


Figura 92. Especificaciones Técnicas de la Jetson Nano [55].

En la Figura 93 se puede visualizar la tarjeta adquirida para realizar el proyecto



Figura 93. Jetson Nano Adquirida

## Anexo C. Implementación del sistema

En la Figura 94 se puede visualizar el prototipo del sistema implementado.

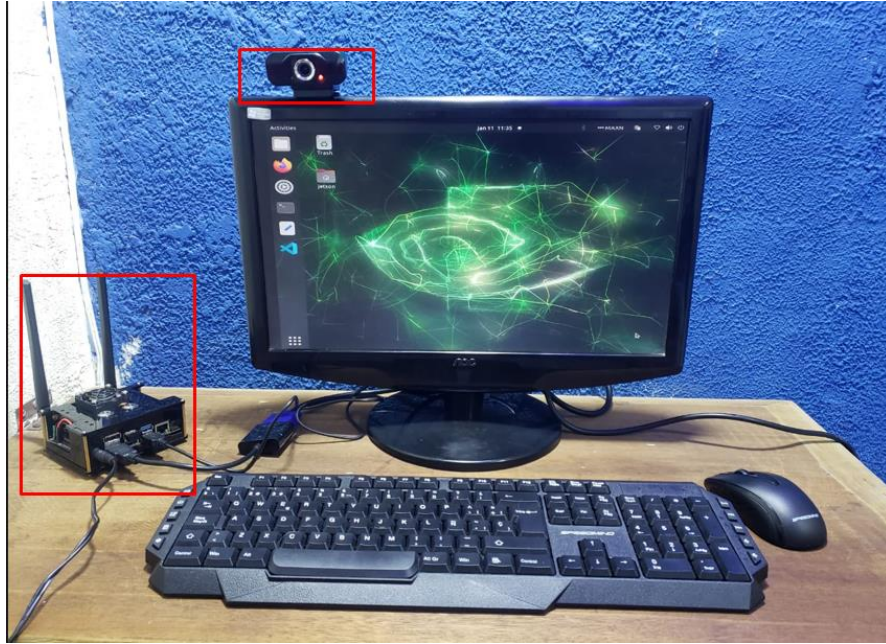


Figura 94. Implementación del sistema.

En la Figura 95 se observa las pruebas en un entorno real del sistema detector de emociones.

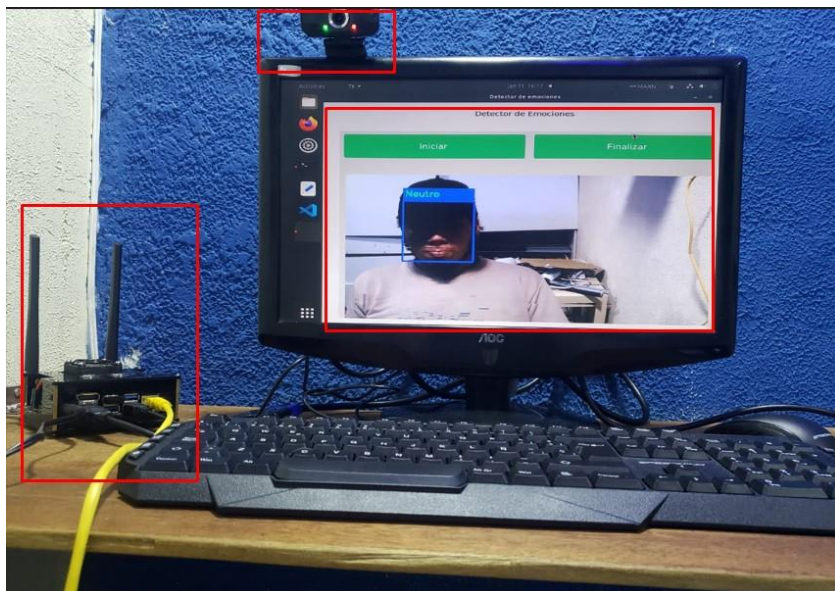


Figura 95. Sistema en un entorno real.

## Anexo D. Implementación del sistema

Programación de toda la interfaz, incluyendo el modelo y clasificación de resultados.

```
#Importamos las librerías necesarias
import cv2
from tensorflow.keras.applications.imagenet_utils import
preprocess_input
from tensorflow.keras.preprocessing.image import img_to_array
from tensorflow.keras.models import load_model
import numpy as np
import imutils
import time
import tkinter as ctk
import tkinter.messagebox as tkmb
from PIL import Image
from PIL import ImageTk
from api import *
from tkinter import ttk
from datetime import datetime
#Definimos variables
usuario=None
examen=None
sesion=None
#Establecemos las clases de nuestra interfaz
def irmenu(elemento):
    global usuario
    usuario=elemento
    view_login.withdraw()
    view_menu.deiconify()

def regresar():
    view_menu.withdraw()
    view_login.deiconify()

def ircamara(elemento):
    global examen
    examen=elemento
    print(elemento)
    view_menu.withdraw()
    view_camera.deiconify()

def regresarmenu():
    global cap
    cap.release()
    view_camera.withdraw()
    view_menu.deiconify()
```

```

#declara view para la cámara
view_camera=ctk.Tk()
view_login = ctk.Tk()

#Asignamos clases para el Login
view_login.geometry("400x400")
view_login.title("Detector de Emociones")

def trylogin():
    resp=postlogin(cedula=user_entry.get(),clave=user_pass.get())
    if(resp['success']):
        irmenu(elemento=resp['data'])
    else:
        tkmb.showerror(title="Login Failed",message="Invalid Username
and password")

label = ctk.Label(view_login,text="Ingreso al sistema de detector de
emociones")
label.pack(pady=20)

#Capturamos los frames
frame = ctk.Frame(master=view_login)
frame.pack(pady=20,padx=40,fill='both',expand=True)

label = ctk.Label(master=frame,text='Detector Emociones')
label.pack(pady=12,padx=10)
user_entry= ctk.Entry(master=frame)
user_entry.pack(pady=12,padx=10)
user_pass= ctk.Entry(master=frame,show="*")
user_pass.pack(pady=12,padx=10)
button = ctk.Button(master=frame,text='Login',command=trylogin)
button.pack(pady=12,padx=10)

#Creamos el Menú

def on_scroll(*args):
    canvas.yview(*args)

view_menu = ctk.Tk()
view_menu.title("Menu")
view_menu.geometry("800x800")

frame = ttk.Frame(view_menu, padding=(3, 3, 12, 12))
frame.grid(column=0, row=0, sticky=(ctk.N, ctk.W, ctk.E, ctk.S))

#Aplicamos canvas para mejorar la interfaz

```

```

canvas = ctk.Canvas(frame, bg="white", width=780, height=780)
canvas.grid(row=0, column=0, sticky=(ctk.N, ctk.W, ctk.E, ctk.S))

scrollbar = ttk.Scrollbar(frame, orient="vertical", command=on_scroll)
scrollbar.grid(row=0, column=1, sticky=(ctk.N, ctk.S))
canvas.config(yscrollcommand=scrollbar.set)

content_frame = ttk.Frame(canvas)
canvas.create_window((0, 0), window=content_frame, anchor=ctk.NW)
content_frame.bind("<Configure>", lambda event, canvas=canvas:
canvas.configure(scrollregion=canvas.bbox("all")))

#Definimos la clase para que se recuerde si el usuario esta activo o no
def on_mousewheel(event):
    canvas.yview_scroll(-1 * (event.delta // 120), "units")

canvas.bind_all("<MouseWheel>", on_mousewheel)
canvas.update_idletasks()
container=ctk.Label(content_frame,text='Lista de Exámenes')
container.pack(padx=10,pady=10,fill="x",expand="yes")
container=ctk.Button(content_frame,text='Cerrar
Sesion',command=regresar)
container.pack(padx=10,pady=15)
examenes=getExamenes()
if(examenes['success']):
    for elemento in examenes['data']:
        if(elemento['ESTADO']=="a"):
            container=ctk.LabelFrame(content_frame,text='Test',width
h=800)

            container.pack(padx=10,pady=10,fill="x",expand="yes")
            label=ctk.Label(container,text="Tema")
            label.pack(padx=20,pady=1)

            label=ctk.Label(container,text=elemento['ASUNTO'])
            label.pack(padx=20,pady=5)
            label1=ctk.Label(container,text="Descripcion:
"+elemento['DESCRIPCION'])
            label1.pack(padx=20,pady=5)
            label2=ctk.Label(container,text="Instrucciones:
"+elemento['INSTRUCCIONES'])
            label2.pack(padx=20,pady=5)
            btn2=ctk.Button(container,text="EMPEZAR",command=lambda
ele=elemento: ircamara(elemento=ele))
            btn2.pack(padx=20,pady=10)
            btn2.pack()
view_menu.withdraw()

```

```

#Activar Camera y capturar frames
emocionActual=None
timeInicio=None
def predict_emotion(frame,faceNet,emotionModel):
    blob = cv2.dnn.blobFromImage(frame, 1.0, (224, 224),(104.0, 177.0,
123.0))
    faceNet.setInput(blob)
    detections = faceNet.forward()
    faces = []
    locs = []
    preds = []
    #Establecemos la condición de entrada del frame
    for i in range(0, detections.shape[2]):
        if detections[0, 0, i, 2] > 0.4:
            box = detections[0, 0, i, 3:7] * np.array([frame.shape[1],
frame.shape[0], frame.shape[1], frame.shape[0]])
            (Xi, Yi, Xf, Yf) = box.astype("int")

            if Xi < 0: Xi = 0
            if Yi < 0: Yi = 0

            face = frame[Yi:Yf, Xi:Xf]
            face = cv2.cvtColor(face, cv2.COLOR_BGR2GRAY)
            face = cv2.resize(face, (48, 48))
            face2 = img_to_array(face)
            face2 = np.expand_dims(face2,axis=0)
            faces.append(face2)
            locs.append((Xi, Yi, Xf, Yf))

            pred = emotionModel.predict(face2)
            preds.append(pred[0])

    return (locs,preds)
def visualizar():
    global emocionActual
    global timeInicio
    #Normalizamos el frame para la detección de rostro
    global cap
    if cap is not None:
        ret,frame=cap.read()
        if ret:
            frame=imutils.resize(frame,width=800,height=600)
            (locs, preds) = predict_emotion(frame,faceNet,emotionModel)
            for(box,pred) in zip (locs,preds):
                (Xi,Yi,Xf,Yf)=box
                (angry,disgust,fear,happy,neutral,sad,surprise) =
pred

                label=''

```



```

        emocionP=format(classes[np.argmax(pred)])
        if(emocionActual==None):
            timeInicio=datetime.now().strftime('%H:%M:%S')
            emocionActual=emocionP

        if(emocionP != emocionActual and emocionActual !=
None):
            guardarDetalleSesion(emocionObj=emocionActual,t
iempoInicio=timeInicio)
            timeInicio=datetime.now().strftime('%H:%M:%S')
            emocionActual=emocionP

        label = "{}".format(classes[np.argmax(pred)])
        cv2.rectangle(frame, (Xi, Yi-40), (Xf, Yi),
(255,0,0), -1)
        cv2.putText(frame, label, (Xi+5, Yi-
15),cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0,255,0), 2)
        cv2.rectangle(frame, (Xi, Yi), (Xf, Yf), (255,0,0),
3)

#Lo transformamos a escala de grises
        frame=cv2.cvtColor(frame,cv2.COLOR_BGR2RGB)
        im=Image.fromarray(frame)
        img=ImageTk.PhotoImage(image=im)
        lblVideo.config(image=img)
        lblVideo.image=img
        lblVideo.after(10,visualizar)
    else:
        lblVideo.image=""
        cap.release()
def iniciar():
    global usuario
    global examen
    global sesion
    global cap
    cap=cv2.VideoCapture(0,cv2.CAP_GSTREAMER)
    sesionObj=postSesion(idExamen=examen['ID'],idUsuario=usuario['ID'])
    if(sesionObj['success']):
        sesion=sesionObj['data'][0]
        visualizar()

def guardarDetalleSesion(emocionObj,tiempoInicio):
    global sesion
    print(sesion)

    hora_actual = datetime.now().strftime('%H:%M:%S')
    postDetalleSesion(idsesion=sesion['ID'],timeInicio=tiempoInicio,ti
meFin=hora_actual,emocionOBJ=emocionObj)

```

```

view_camera.title("Detector Facial")
view_camera.geometry("800x500")
cap=None

#declarar el modelo
time_actualframe = 0
time_prevframe = 0

# Tipos de emociones del detector
#classes =
['angry','disgust','fear','happy','neutral','sad','surprise']
classes =
['enojado','Asco','Miedo','Feliz','Neutro','Triste','Sorprendido']

# Cargamos el modelo de detección de rostros
prototxtPath = "face_detector/deploy.prototxt"
weightsPath = "face_detector/res10_300x300_ssd_iter_140000.caffemodel"
faceNet = cv2.dnn.readNet(prototxtPath, weightsPath)
# Carga el detector de clasificación de emociones
emotionModel = load_model("model.h5")
#iniciar modelo
imaget_path = "init.jpg"
faces=[]
# Redimensionamos la imagen y convertimos a gray
face = cv2.cvtColor(cv2.imread(imaget_path), cv2.COLOR_BGR2GRAY)
face = cv2.resize(face, (48, 48))
face2 = img_to_array(face)
face2 = np.expand_dims(face2,axis=0)
faces.append(face2)
# El modelo estima la predicción
preds = emotionModel.predict(faces)
print(classes[np.argmax(preds)])

btnIniciar=ctk.Button(view_camera,text='Iniciar',width=45,command=ini
ar)
btnIniciar.grid(column=0,row=0,padx=5,pady=5)
btnFinalizar=ctk.Button(view_camera,text='Finalizar',width=45,command=r
egresarmenu)
btnFinalizar.grid(column=1,row=0,padx=5,pady=5)
lblVideo=ctk.Label(view_camera)
lblVideo.grid(column=0,row=1,columnspan=2)
view_camera.withdraw()
view_login.mainloop()
#view_menu.mainloop()

```



## Programación para el API con en web Hosting

```
import requests
import json
from datetime import datetime
base_url = "https://detector-emociones.000webhostapp.com/APIFACE/API/"

def postlogin(cedula, clave):
    url = base_url + "login.php"

    auth = {
        'cedula':cedula,
        'clave': clave
    }

    headers = {
        'Accep':'application/json',
        'Content-Type': 'application/json'
    }

    response = requests.request("POST", url,
headers=headers,data=json.dumps(auth))
    data=response.content
    obj=json.loads(data)
    print(obj)
    return obj

def getExamenes():
    url = base_url + "examen.php"

    headers = {
        'Accep':'application/json',
        'Content-Type': 'application/json'
    }

    response = requests.request("GET", url, headers=headers)
    data=response.content
    obj=json.loads(data)
    return obj

def postDetalleSesion(idsesion,timeInicio,timeFin,emocionOBJ):
    url = base_url + "detalleseesion.php"
```

```

auth = {
    'ID': 0,
    'ID_SESION_PER': idsesion,
    'TIEMPO_INI': timeInicio,
    'TIEMPO_FIN': timeFin,
    'EMOCION': emocionOBJ
}

headers = {
    'Accep': 'application/json',
    'Content-Type': 'application/json'
}

response = requests.request("POST", url,
headers=headers,data=json.dumps(auth))
data=response.content
obj=json.loads(data)
return obj
def postSesion(idExamen,idUsuario):
url = base_url + "sesion.php"
fecha_actual = datetime.now()
formato_deseado = fecha_actual.strftime('%Y-%m-%d %H:%M:%S')
auth = {
    'ID': '0',
    'ID_EXAMEN_PER': idExamen,
    'ID_USUARIO': idUsuario,
    'FECHA': formato_deseado,
    'OBSERVACIONES': 'PRUEBA',
    'ESTADO': 'a'
}

headers = {
    'Accep': 'application/json',
    'Content-Type': 'application/json'
}

response = requests.request("POST", url,
headers=headers,data=json.dumps(auth))
data=response.content
obj=json.loads(data)
return obj

```