**UNIVERSIDAD NACIONAL DE INGENIERÍA**

**UNIDAD DE POSGRADO FACULTAD DE INGENIERÍA INDUSTRIAL Y SISTEMAS / ESCUELA DE POSGRADO**



**TESIS**

**“MEDICIÓN AUTOMÁTICA DE ESTADOS AFECTIVOS DE ESTUDIANTES UNIVERSITARIOS EN AULAS HÍBRIDAS: UNA SOLUCIÓN EDGE AI OPTIMIZADA MEDIANTE ANÁLISIS MULTIMODAL (FACIAL Y POSTURAL)”**

PARA OBTENER EL GRADO ACADÉMICO DE MAESTRO CON MENCIÓN EN:  
MAESTRÍA EN INTELIGENCIA ARTIFICIAL

**ELABORADA POR:**

**GEYSON DAVID MORALES CCASA**

**JOSE LUIS ESPINOZA GARCIA**

**ASESOR:**

**SAMUEL OPORTO DIAZ**

**LIMA, PERÚ**

**2025**

**Índice**

Contenido

[Resumen – Abstract 5](#_Toc213605854)

[Introducción 6](#_Toc213605855)

[CAPÍTULO i. PLANTEAMIENTO DEL PROBLEMA 7](#_Toc213605856)

[1.1. Diagnóstico 7](#_Toc213605857)

[1.2. Identificación y Descripción del Problema de Estudio 7](#_Toc213605858)

[1.3. Formulación del Problema 8](#_Toc213605859)

[1.3.1. Formulación del Problema General 8](#_Toc213605860)

[1.3.2. Formulación de los Problemas Específicos 9](#_Toc213605861)

[1.4. Objetivos 9](#_Toc213605862)

[1.4.1. Objetivo General 9](#_Toc213605863)

[1.4.2. Objetivos Específicos 9](#_Toc213605864)

[3.1. Metodología 10](#_Toc213605865)

[CAPÍTULO II: MARCO TEÓRICO 13](#_Toc213605866)

[4.1. Antecedentes bibliográficos 13](#_Toc213605868)

[4.2. Bases Teóricas 14](#_Toc213605869)

[4.3. Definición de Términos 15](#_Toc213605870)

[CAPÍTULO III: PROPUESTA Y DESARROLLO DE LA SOLUCIÓN 15](#_Toc213605871)

[5.1. Presentación de la Propuesta de Solución 15](#_Toc213605873)

[5.2. Desarrollo de la Propuesta de Solución 16](#_Toc213605874)

[CAPÍTULO IV. CONCLUSIONES Y RECOMENDACIONES 19](#_Toc213605875)

[6.1. CONCLUSIONES 19](#_Toc213605877)

[6.2. RECOMENDACIONES 19](#_Toc213605878)

[REFERENCIAS 19](#_Toc213605879)

[ANEXO 20](#_Toc213605880)

[ANEXO 1. Resumen de Antecedentes y Artículos de investigación 20](#_Toc213605881)

[ANEXO 2. Escala Likert 21](#_Toc213605882)

Lista de Ilustraciones

[Ilustración 1. Metodología de la Investigación del proyecto de investigación. 10](#_Toc213602199)

[Ilustración 2. Proceso de Construcción del artefacto para el proyecto. 13](#_Toc213602200)

Lista de Tablas y Cuadros

[Tabla 1. Listado de objetivos específicos. 8](#_Toc213602201)

[Tabla 3. Fases metológicas para desarrollo del prototipo. 8](#_Toc213602202)

[Tabla 2. Listado de variables dependientes e independientes del proyecto. 10](#_Toc213602203)

[Tabla 4. Componentes de la arquitectura de solución modular. 12](#_Toc213602204)

[Tabla 5. Listado resumido de los antecedentes y fuentes bibliográficas. 16](#_Toc213602205)

[Tabla 6. Matriz de Consistencia 1](#_Toc213602206)

## Resumen – Abstract

## Introducción

El panorama educativo ha evolucionado hacia el modelo de aulas híbridas, donde la experiencia de aprendizaje combina la instrucción presencial con la participación remota. Si bien este modelo promueve la flexibilidad, presenta un desafío fundamental: la dificultad del docente para monitorear eficazmente el nivel de atención y el compromiso o atención de los estudiantes en tiempo real, especialmente aquellos conectados a distancia.

La atención es un indicador clave del éxito académico, pero las técnicas de observación tradicionales se vuelven insuficientes ante la asimetría de datos entre el aula física y el entorno virtual. Las soluciones existentes basadas en Inteligencia Artificial (IA) suelen depender de la computación en la nube, lo que introduce problemas de latencia inaceptable para una retroalimentación en tiempo real y preocupaciones éticas relativas a la privacidad de los datos sensibles de los estudiantes.

Esta tesis propone abordar esta brecha mediante el desarrollo de un sistema de Edge AI (Inteligencia Artificial en el Borde). Este sistema se enfoca en la optimización de modelos ligeros de Machine Learning y Deep Learning para que puedan ejecutar inferencias de análisis facial y postural directamente en dispositivos menores, superando las limitaciones de la dependencia de la nube y ofreciendo una herramienta de monitoreo y asistencia en tiempo real de alta usabilidad y eficiencia.

## CAPÍTULO i. PLANTEAMIENTO DEL PROBLEMA

# Diagnóstico

La transición al modelo de aulas híbridas, impulsada por la necesidad de flexibilidad, ha expuesto una **brecha de monitoreo y retroalimentación** en tiempo real. La eficacia del proceso de enseñanza-aprendizaje está intrínsecamente ligada al **nivel de atención** y participación del estudiante. En el aula tradicional, el docente emplea la observación directa de señales no verbales (gestos, postura, contacto visual) para medir el compromiso o atención. En el entorno híbrido, esta observación es severamente limitada:

1. **Distancia Visual:** Es casi imposible monitorear simultáneamente a los estudiantes presenciales y a los remotos, cuyas cámaras a menudo están apagadas o solo muestran una pequeña parte de su rostro.
2. **Asimetría de Datos:** Los estudiantes remotos y en presencial no comparten el mismo espacio físico, lo que oculta señales conductuales no verbales cruciales. El documento plantea claramente: "los docentes enfrentan dificultades para monitorear el nivel de atención de los estudiantes en tiempo real, especialmente cuando estos se conectan de manera remota.
3. **Latencia en la Retroalimentación:** Los sistemas tradicionales de analíticas de aprendizaje a menudo procesan datos a posteriori, lo que anula la posibilidad de una intervención pedagógica *in situ* para reenganchar al estudiante.
4. **Restricciones de Infraestructura:** En muchas instituciones educativas, el ancho de banda y los recursos de cómputo son limitados, haciendo inviable el uso de soluciones de IA que requieren alta capacidad de procesamiento en la nube.

# Identificación y Descripción del Problema de Estudio

En los entornos de aulas híbridas universitarias, donde la interacción entre estudiantes presenciales y remotos se realiza de manera simultánea, se ha evidenciado una disminución en la atención, motivación y participación activa de los alumnos debido a la falta de mecanismos que permitan evaluar en tiempo real su estado afectivo.  
Los docentes carecen de herramientas objetivas que les ayuden a detectar indicadores de distracción, fatiga o interés, lo que dificulta la adaptación inmediata de las estrategias pedagógicas.

Aunque existen sistemas de análisis de emociones basados en visión artificial, la mayoría requieren infraestructura en la nube con alta latencia y dependencia de conexión, lo que los hace poco viables para aulas con limitaciones tecnológicas o de conectividad. Además, no integran múltiples modalidades de análisis (facial y postural) que brinden una visión más precisa del estado afectivo del estudiante.

# Formulación del Problema

Según (Hossen & Uddin, 2023), las soluciones de Inteligencia Artificial (IA) existentes para el monitoreo de atención suelen basarse en la **computación en la nube**. Si bien son potentes, introducen serios problemas que limitan su aplicabilidad en contextos educativos reales, especialmente en regiones con infraestructura limitada.

# Formulación del Problema General

El problema central radica en la falta de accesibilidad a una herramienta, ligera que permita medir en tiempo real los estados afectivos de los estudiantes en aulas híbridas, combinando información facial y postural mediante Edge AI, garantizando bajo consumo computacional y alta precisión.

El presente proyecto de investigación abordará la **"****Medición Automática de Estados Afectivos de Estudiantes Universitarios en Aulas Híbridas”.** Por medio de una **“Solución Edge AI Optimizada Mediante Análisis Multimodal (Facial y Postural)".**

Se basa en el concepto de "Medición Automática de Estados Afectivos" donde se establece una medición para este tipo de escenarios de atención para estudiantes (the Central China Normal University, China et al., 2019).

# Formulación de los Problemas Específicos

Entre los principales problemas tenemos:

* **Alta Latencia:** El envío de *streams* de video al *cloud* para procesamiento y la espera de la respuesta generan una latencia que hace que la retroalimentación al docente no sea verdaderamente en **tiempo real**, frustrando la intervención pedagógica inmediata.
* **Consumo de Ancho de Banda:** El tráfico constante de video de múltiples cámaras consume una gran cantidad de ancho de banda, lo que resulta costoso e inviable en entornos con conectividad limitada.
* **Riesgos de Privacidad:** El procesamiento de datos sensibles (análisis facial, biometría emocional) fuera de las instalaciones educativas (en servidores de terceros en la nube) incrementa los riesgos de seguridad y cumplimiento normativo (ej. Ley N° 29733 en Perú).
* **La falta de accesibilidad de Equipamiento:** La dependencia de servidores *cloud* implica un gasto operativo recurrente elevado, inaccesible para muchas instituciones educativas.

# Objetivos

# Objetivo General

Validar un sistema Edge AI optimizado para la medición automática y en tiempo real de los estados afectivos (atención, distracción, fatiga) de estudiantes universitarios en aulas híbridas, logrando una alta eficiencia computacional y baja latencia.

# Objetivos Específicos

Tenemos los siguientes:

1. Identificar, seleccionar y extraer las características faciales (emociones, mirada, expresión neutra o distraída) y posturales (posición de cabeza, inclinación corporal) más representativas que se asocian directamente con los estados de atención, distracción y fatiga en estudiantes universitarios.
2. Implementar y evaluar técnicas de optimización (ej. cuantización de 8 bits, poda) en los modelos ligeros de detección para reducir la carga computacional y la latencia, asegurando un rendimiento óptimo en dispositivos Edge.
3. Evaluar el desempeño del prototipo funcional de Edge AI en condiciones reales de aula híbrida, validando su precisión, latencia y la utilidad pedagógica percibida.

Tabla . Listado de objetivos específicos.

|  |  |  |  |
| --- | --- | --- | --- |
| **N°** | **Objetivo Específico** | **Métrica/Indicador Clave** | **Descripción y Justificación** |
| **OE 1** | **Adquisición y Normalización de Datos Afectivos Multimodales** | Tasa de Muestras Útiles (≥95% de *frames* anotados correctamente). | **Enfocado en la Brecha de *Dataset***. Consiste en la conceptualización, creación y anotación de un *dataset* local de video (facial y postural) en un entorno de aula híbrida simulado/real, vital para el entrenamiento y *fine-tuning* de modelos especializados para la población universitaria. |
| **OE 2** | **Optimización y Despliegue de Modelos Ligeros en Arquitectura Edge** | Precisión del Modelo (≥90%),  Tasa de FPS (≥25), Reducción de Latencia (≥50% vs. *Cloud-based*). | **Enfocado en la Solución Tecnológica**. Implementar arquitecturas ligeras (ej. MobileNet, YOLO-Nano) y aplicar técnicas de optimización (cuantización a 8-bit, poda) para alcanzar la meta de **tiempo real** en dispositivos de borde (ej. NVIDIA Jetson Nano o Raspberry Pi 5). |
| **OE 3** | **Diseño e Integración del Prototipo Funcional y *Dashboard*** | Usabilidad (≥4.0/5.0 en escala Likert), Precisión de Alerta (≥80%). | **Enfocado en el Impacto Aplicado**. Desarrollar la lógica de inferencia, la interfaz de alertas en tiempo real y la visualización histórica (*dashboard*) para el docente, validando su **utilidad pedagógica** y **usabilidad**. |

# Metodología

Se propone una metodología de **Investigación** con un enfoque iterativo, siguiendo un proceso de Ingeniería de *Machine Learning* para *Edge AI*.

El proyecto se abordará con un enfoque de **Investigación de Diseño** (*Design Science Research - DSR*), dado que el objetivo principal es la creación de un artefacto tecnológico (el sistema Edge AI). La validación del artefacto tendrá un **enfoque cuantitativo** para medir la precisión, el rendimiento técnico y la usabilidad.

* **Tipo de Investigación:** Aplicada, con énfasis en el desarrollo tecnológico y la validación empírica.
* **Diseño de Investigación:** No Experimental. La investigación se centrará en el desarrollo y la evaluación de un prototipo.
* **Fases del DSR:** Se seguirán las fases de DSR: 1) Definición del Problema y Justificación; 2) Definición de los Objetivos del Artefacto; 3) Diseño y Desarrollo; 4) Demostración (Prototipo funcional); y 5) Evaluación (Validación de precisión y usabilidad)

Tabla . Fases metodológicas para desarrollo del prototipo.

|  |  |  |
| --- | --- | --- |
| Fase Metodológica | Descripción y Tarea Clave | Consistencia con OE |
| Fase I: Conceptualización y Adquisición de Datos | **Revisión Sistemática:**  Identificación de las mejores arquitecturas (*MobileNet, YOLO-Nano, Mini-Xception*) y técnicas de optimización.  **Creación/Elección del *Dataset* (OE 1):**  Definición del protocolo de captura y anotación (etiquetado manual) de videos de estudiantes (universitarios, aula híbrida simulada) bajo las clases de **Atención, Distracción y Fatiga**. | OE 1 |
| Fase II: Optimización y Entrenamiento *Edge* | **Selección de Modelo y *Fine-Tuning***:  Uso de un modelo preentrenado (ej. Mini-Xception para emoción, OpenPose/MoveNet para postura) y adaptación al *dataset* local.  **Optimización (OE 2):** Aplicación de **Cuantización Post-Entrenamiento (PTQ)** y evaluación de la reducción de la carga computacional en el dispositivo de borde. | OE 2 |
| Fase III: Desarrollo e Integración del Prototipo | **Integración *Edge*:**  Despliegue del modelo optimizado en la plataforma de borde (ej. Jetson Nano) utilizando *frameworks* de inferencia rápida (*TensorFlow Lite, OpenVINO*).  **Elaboración del *Dashboard* (OE 3):**  Construcción de la interfaz de visualización de datos en tiempo real (alertas) e históricos para el docente. | OE 2, OE 3 |
| Fase IV: Validación y Evaluación Aplicada | **Prueba en Escenario Real (OE 3):**  Medición de la Latencia (ms) y FPS en condiciones reales.  **Validación de Impacto:**  Aplicación de la **Encuesta de Usabilidad (SUS/Likert)** a docentes para evaluar la utilidad pedagógica del sistema de alertas. | OE 3 |

**Diagrama

El contenido generado por IA puede ser incorrecto.**

Ilustración . Metodología de la Investigación del proyecto de investigación.

**Variables del Proyecto**

Se establecen las siguientes variables para el proyecto.

Tabla . Listado de variables dependientes e independientes del proyecto.

|  |  |  |  |
| --- | --- | --- | --- |
| Tipo de Variable | Denominación | Indicadores / Clases / Niveles | Objetivo Específico Asociado |
| Variable Independiente (VI​) | **Sistema Edge AI Optimizado** | 1. Arquitectura ML/DL y Técnica de Optimización (Cuantización, Poda).  2. Rendimiento en FPS, Latencia, Uso de CPU. | OE 2 |
| Variable Dependiente (VD​) | **Medición de Estados Afectivos** (proxy de Atención) | Clases de Estados:  Atención, Distracción (mirada fuera), Fatiga (parpadeo, postura encorvada).  Métrica:  Precisión (*Accuracy*),  F1-*Score* de Clasificación. | OE 1, OE 2 |
| Variables de Control (Parámetros) | **Contexto de Aplicación** | 1. Entorno: Aulas Híbridas (Presencial y Remoto).  2. Población: Estudiantes de Nivel Universitario.  3. Fuente de Datos: Análisis Multimodal (Facial y Postural). | Global (OE 1, OE 2, OE 3) |

## CAPÍTULO II: MARCO TEÓRICO



# Antecedentes bibliográficos

La revisión bibliográfica confirma la relevancia de la detección de estados afectivos y la necesidad de soluciones eficaces.

1. **Detección de Engagement y Estados Afectivos**:
   * **Hasnine et al. (2021, 2023):** Demuestran la efectividad de la extracción y visualización de emociones para la detección de engagement en el aprendizaje en línea, centrándose en el análisis de video. Su trabajo justifica la base del análisis facial, aunque no se enfoca en Edge AI.
   * **Hossen et al. (2023):** Utilizan clasificadores XGBoost para el monitoreo de la atención en clases en línea, destacando la importancia del análisis no verbal. Esto apoya la selección de modelos y la relevancia de las características.
   * **Wang et al. (2019):** Miden estados afectivos (como engagement, frustración, confusión) a través de registros de aprendizaje en línea, lo que proporciona una base conceptual para la categorización de los estados de atención.
2. **Edge Computing y Optimización de Modelos:** 
   * **Abdulkader et al. (2023):** Abordan la optimización del engagement en el aprendizaje en línea basado en el borde (edge-based), validando el enfoque de llevar el procesamiento de datos al dispositivo local para mejorar la eficiencia y la privacidad.
   * **Gao et al. (2021) / Pang et al. (2020):** Sustentan la viabilidad técnica y la necesidad de modelos optimizados (como YOLOv4/v7 para postura y rostro) mediante técnicas como la cuantización y la poda (pruning) para su ejecución en dispositivos de bajo costo (TinyML/Edge Computing).

# Bases Teóricas

* + 1. **Edge AI (Inteligencia Artificial en el Borde)**

Se define como la ejecución de algoritmos de IA directamente en un dispositivo local (*edge device*) en lugar de en la nube. Esto minimiza la latencia, reduce los costos de transmisión de datos, disminuye el consumo de energía y mejora la privacidad. Es el pilar de la solución propuesta para el monitoreo en tiempo real.

* + 1. **Monitoreo de Atención Multimodal**

La detección de la atención se realizará mediante la combinación de dos modalidades visuales:

* **Análisis Facial**: Detección de emociones (ej. neutro, concentración vs. aburrimiento, frustración), detección de puntos clave (land-marks) y análisis de la mirada (dirección, parpadeo) para inferir el nivel de *engagement*.
* **Análisis Postural (Pose Estimation):** Uso de modelos ligeros para detectar la posición de la cabeza, la inclinación del cuerpo y la presencia/ausencia del estudiante, lo cual es indicativo de distracción o fatiga.
  + 1. **Optimización de Modelos Ligeros**

Para garantizar la eficiencia en el Edge, se emplearán técnicas avanzadas:

* **Cuantización:** Reducción de la precisión numérica de los parámetros del modelo (ej. de *float32* a *int8*), lo que disminuye el tamaño del modelo y acelera la inferencia sin una pérdida significativa de precisión.
* **Poda (*Pruning*):** Eliminación de conexiones o neuronas redundantes en la red, resultando en un modelo más pequeño y rápido.

# Definición de Términos

* **Edge AI**: Inteligencia Artificial procesada localmente en el dispositivo de captura de datos (borde de la red), sin depender del servidor en la nube.
* **Aula Híbrida:** Modelo educativo que combina la instrucción presencial con la participación remota de estudiantes.
* **Latencia:** Tiempo de retardo entre la captura de los datos (video) y la generación de la inferencia (alerta de atención).
* **Modelo Ligero:** Modelo de Machine Learning/Deep Learning que ha sido optimizado para tener un tamaño reducido y un bajo consumo computacional, ideal para Edge Devices.
* **Multimodalidad:** Uso de diferentes tipos de datos (facial, postural, en este caso) para la toma de decisiones o la inferencia.
* **Cuantización:** Proceso de optimización que reduce la precisión numérica de los pesos de un modelo de IA para acelerar su ejecución.
* **Engagement:** Implicación activa, compromiso y participación del estudiante en el proceso de aprendizaje.

## CAPÍTULO III: PROPUESTA Y DESARROLLO DE LA SOLUCIÓN



# Presentación de la Propuesta de Solución

La solución se centra en **Edge AI** (IA en el borde), que traslada el procesamiento y la inferencia a dispositivos locales de bajo costo (ej. microcontroladores, Raspberry Pi, Jetson Nano) cerca de la fuente de datos. Esto es la justificación principal para tu pregunta general: **"¿Cómo aplicar Edge AI para detectar en tiempo real...?"** *[Matriz de Consistencia]*.

**Al implementar Edge AI, se logra:**

* **Baja Latencia y Detección en Tiempo Real:** Las inferencias se ejecutan localmente, (Por ejemplo garantizando una tasa de *frames* por segundo (FPS) superior a 25 y una retroalimentación casi instantánea)
* **ptimización y Eficiencia:** Se requiere la **optimización de modelos ligeros** (mediante cuantización, poda o arquitecturas como MobileNet o EfficientNet) para reducir la carga computacional en el dispositivo de borde, como lo indican tus objetivos específicos.
* **Privacidad:** Los datos biométricos se procesan localmente, minimizando el riesgo de transferencias de datos sensibles a terceros (una consideración clave).

# Desarrollo de la Propuesta de Solución

El progreso inicial del proyecto debe demostrar un avance tangible hacia el **Objetivo Específico 1 (Adquisición y Normalización de Datos)**, que es la base para el entrenamiento.

* + 1. **Herramientas utilizadas para la Construcción**

Todas las herramientas necesarias para la ingesta de datos, preprocesamiento y experimentación inicial están operativas.

* **Entorno de Desarrollo:** Python (v3.9+), PIP, Conda.
* **Frameworks de Desarrollo y *Fine-Tuning***: **TensorFlow/Keras** (para modelado) y **PyTorch**.
* **Frameworks de Visión por Computadora (OE 1):** **OpenCV** (para captura y preprocesamiento de *frames*), **Dlib** o **Mediapipe** (para detección de *landmarks* faciales y esqueleto corporal).
* **Frameworks de Optimización (*Target Edge*) (OE 2):** **TensorFlow Lite (TFLite)** o **OpenVINO** (Intel).
  + 1. **Diseño Diagramado de Componentes y Clases Requeridas**

Se propone una arquitectura modular de tres capas, ideal para la lógica de Edge AI.

Tabla . Componentes de la arquitectura de solución modular.

|  |  |  |
| --- | --- | --- |
| Componente | Capa (Módulo) | Descripción de Clase o Función Clave |
| Módulo de Captura y Preprocesamiento | Ingesta de Datos (*Edge*) | Clase CameraStreamManager: Gestiona la captura de video y la sincronización de *frames* con la anotación de tiempo. |
| Módulo de Detección Multimodal | Inferencias (*Edge*) | Clase FaceDetector: Implementa el modelo ligero (ej. Mini-Xception) para clasificación de emociones (proxy de atención/fatiga). Clase PoseEstimator: Implementa el modelo ligero (ej. MoveNet) para el ángulo de la cabeza/cuerpo. |
| Módulo de Fusión de Datos y Lógica de Alerta | Lógica de Negocio (*Edge*) | Clase AttentionFusionEngine: Combina los *outputs* de rostro y postura para clasificar el **Estado Afectivo** (Atención → 1, Distracción → 0). Función AlertGenerator: Genera la señal de alerta si el estado persiste por 'N' segundos. |
| Módulo de Visualización | Interfaz (*Local/Web*) | Clase DashboardRenderer: Muestra la *stream* de video con los *bounding boxes* y la clasificación del estado afectivo, junto al historial de alertas para el docente. |

**Diagrama

El contenido generado por IA puede ser incorrecto.**

Ilustración . Proceso de Construcción del artefacto para el proyecto.

* + 1. **Componentes Construidos para el Objetivo 1**

El avance del **OE 1 (Adquisición de Datos)** se demuestra con la construcción del **Procedimiento de Adquisición de Datos** y el mecanismo de **Visualización de Datos** para la anotación.

**A. Procedimiento de Adquisición de Datos (Protocolo de Etiquetado)**

Se ha diseñado un protocolo para la creación del *dataset* inicial:

1. **Sujetos:** N estudiantes universitarios voluntarios.
2. **Escenario:** Sesión de clase de 30 minutos (simulando aula híbrida).
3. **Etiquetas de Verdad Fundamental:**

* **Atención:** Mirada fija a la pantalla, postura erguida.
* **Distracción:** Mirada fuera de pantalla (≥2 segundos), manipulación de objetos no relacionados.
* **Fatiga:** Bostezo, frotarse los ojos, cabeza reclinada sobre el escritorio o inclinación lateral.

1. **Método de Anotación:** Grabación de video. Se utiliza una herramienta de anotación (*Labelling Tool*) donde un anotador humano (experto pedagogo) revisa el video y etiqueta los segmentos de 5 segundos con la clase afectiva predominante.

**B. Mecanismos de Visualización de Datos**

Se ha desarrollado un *script* inicial en Python/OpenCV que demuestra la capacidad de **capturar, preprocesar y visualizar** los *landmarks* faciales y posturales, lo que sirve como la base de la herramienta de anotación (OE 1).

* **Artefacto Elaborado (OE 1):**

|  |  |
| --- | --- |
| Elemento | Descripción para la Demostración |
| Stream de Video | Captura de video en tiempo real, mostrando la vista que tendrá el Edge Device. |
| Detección Facial | Implementación de una librería rápida (ej. MTCNN/MediaPipe Face) para dibujar el *bounding box* alrededor del rostro. |
| Extracción de *Landmarks* | Visualización de los 68 puntos (o más) de *landmarks* faciales y los puntos clave del esqueleto corporal (hombros, cuello) sobre el *stream*. |
| Mecanismo de Etiquetado (*Mock*) | Un botón o *hotkey* que simula la función de un anotador para guardar el *frame* con una etiqueta (ej. "Guardar\_Distracción"). |

## CAPÍTULO IV. CONCLUSIONES Y RECOMENDACIONES



# CONCLUSIONES

# RECOMENDACIONES

## REFERENCIAS

## ANEXO

## ANEXO 1. Resumen de Antecedentes y Artículos de investigación

Tabla . Listado resumido de los antecedentes y fuentes bibliográficas.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Autor(es) y Año | Título / Tema Principal | Metodología / Enfoque | Contribución Relevante para la Tesis | Gaps Identificados / Limitaciones |
| Hasnine et al. (2023) [5] | Detección de Estados Afectivos y Monitoreo de Atención en Línea | MTCNN, Mini-Xception, Dashboard de análisis de aprendizaje en tiempo real. Clasificación de *engagement* y concentración. | Base para arquitectura de monitoreo en tiempo real y visualización de datos de participación. | No enfocado en optimización para Edge AI en dispositivos de bajo costo. |
| Abdulkader et al. (2023) [3] | Optimización del Compromiso Estudiantil en Aprendizaje en Línea Basado en Edge | EBSAAS (Edge-based Student Attentiveness Analysis System). MTCNN, DLIB (EAR, MAR), VGG-19. | Referente directo para Edge AI en monitoreo de atención y uso de características faciales. | Limitado a características faciales; no integra otras modalidades de forma exhaustiva. |
| Wang et al. (2019) [6] | Medición de Estados Afectivos a través de Logs de Aprendizaje en Línea | Análisis de logs de LMS. | Base para integrar datos textuales y de interacción con plataformas de aprendizaje. | No aplica Edge AI; no considera datos visuales o de voz. |
| Hossen & Uddin (2023) [2] | Monitoreo de Atención de Estudiantes en Clases en Línea usando Clasificador XGBoost | Detección facial (MTCNN), extracción de características (DLIB), clasificación (XGBoost). | Demuestra alta precisión con XGBoost para atención; valida el uso de características faciales. | No explora la optimización de modelos para el despliegue en el borde de forma detallada. |
| Sepahvand et al. (2023) [7] | Algoritmo de Aprendizaje Adaptativo Teacher-Student con Destilación de Conocimiento para Inteligencia en el Borde | Destilación de conocimiento (KD) para modelos ligeros. | Crucial para la "Optimización de Modelos Ligeros" en dispositivos de borde. | No se enfoca en el dominio educativo ni en el análisis multimodal específico. |
| Boiko et al. (2024) [8] | Arquitecturas Edge-Cloud para Sistemas Híbridos de Gestión Energética: Una Revisión Exhaustiva | Revisión de arquitecturas *edge-cloud*. | Visión general de arquitecturas *edge-cloud* y métricas de rendimiento para la distribución de carga computacional. | No aplica directamente al dominio educativo; es una revisión conceptual. |
| Pang et al. (2023) [4] | Red Convolucional de Grafos para la Detección Automática del Comportamiento No Verbal de Docentes | Faster R-CNN, OpenPose (puntos esqueléticos), GCN. | Relevante para el análisis postural y tipificación de comportamientos no verbales en el aula. | Se centra en docentes, no en estudiantes; no explora Edge AI ni otras modalidades. |

**Fuente:** Elaboración Propia.

## ANEXO 2. Escala Likert

**¿Qué es la escala Likert?**

La escala de Likert es una escala psicométrica de calificación que se utiliza para medir el nivel de acuerdo o desacuerdo de una persona con una declaración o afirmación específica. Fue desarrollada en 1932 por el psicólogo estadounidense Rensis Likert como un instrumento para medir opiniones, actitudes y comportamientos.​

**Características principales**

A diferencia de las preguntas simples de "sí" o "no", la escala de Likert permite a los encuestados calificar sus respuestas en un rango de opciones que va desde un extremo a otro, capturando así la intensidad de sus sentimientos hacia un tema determinado. Esta escala asume que la fuerza e intensidad de la experiencia es lineal, yendo típicamente desde "totalmente de acuerdo" hasta "totalmente en desacuerdo".​

**Formato de respuestas**

Las opciones de respuesta más comunes incluyen:​

* Completamente de acuerdo
* De acuerdo
* Neutral (ni de acuerdo ni en desacuerdo)
* En desacuerdo
* Completamente en desacuerdo

Las escalas pueden configurarse con diferentes niveles de medición, siendo las más utilizadas las de 5, 7 y 9 puntos. Se ha demostrado que la escala de 7 puntos alcanza los límites superiores de confiabilidad, añadiendo "muy" a los extremos de la escala de 5 puntos.​

**Aplicaciones**

La escala de Likert es ampliamente utilizada en investigación de mercados, ciencias sociales, psicología, recursos humanos y negocios para detectar actitudes, opiniones e intenciones de consumidores o empleados. Es particularmente útil para obtener feedback preciso sobre productos, servicios o acciones empresariales, permitiendo determinar si algo se percibe como "suficientemente bueno" o "excelente", o si requiere ajustes.

Tabla . Matriz de Consistencia

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **MATRIZ DE CONSISTENCIA** | | | | | |
| **Título:** Medición Automática en Tiempo Real de Estados Afectivos de Estudiantes Universitarios en Aulas Híbridas: Una Solución Edge AI Optimizada Mediante Análisis Multimodal (Facial y Postural) | | | | | |
| **Problema de Estudio** | **Problemas** | **Objetivos** | **Hipótesis** | **Variables y Dimensiones** | **Diseño Metodológico** |
| En las aulas híbridas, se encuentra la dificultad para monitorear el nivel de atención de los estudiantes en tiempo real, especialmente en ambientes híbridos (remoto y presencial).  Las soluciones actuales basadas en inteligencia artificial dependen de la computación en la nube, lo que genera alta latencia, riesgos de privacidad, además de modelos complejos de visión por computador. | **Problema general**  Existe una dificultad clave para el monitoreo en tiempo real del nivel de atención y participación de los estudiantes en aulas híbridas (presenciales y remotas). Las soluciones de IA actuales realizan detección de gestos y posturas identificando patrones, pero al depender de la computación en la nube, generan alta latencia, consumo de recursos computacionales e incrementan de riesgos de privacidad, lo que limita su aplicabilidad práctica en entornos educativos. | **Objetivo general**  Diseñar y validar una solución Edge AI que identifique en tiempo real el nivel de atención estudiantil en aulas híbridas mediante el análisis facial y postural, optimizando modelos ligeros de visión por computador para dispositivos con recursos computacionales limitados. | **Hipótesis general**  La implementación de un sistema de Edge AI basado en análisis facial y postural optimizado con modelos ligeros permite detectar con alta precisión y baja latencia el nivel de atención estudiantil en aulas híbridas, mejorando la capacidad de retroalimentación docente. | **Variable Independiente:**  Sistema Edge AI Optimizado (Modelo) para Detección de Atención. (Ejm: XGBoost)  **Dimensiones:**   * Modelos (Cuantización, Poda), * Procesamiento Multimodal (Análisis Facial y Postural) * Despliegue en Borde. | **Tipo de Investigación:**  Aplicada (Desarrollo de un artefacto tecnológico).  **Nivel de Investigación:**  Descriptivo, Explicativo y Tecnológico.  **Métodos:**  Revisión bibliográfica, Modelado/Simulación, Desarrollo de Prototipo, Evaluación Empírica  **Diseño:**  No Experimental - Investigación de Diseño (Design Science Research) para el desarrollo del artefacto. |
| **Problemas específicos**  1. Deficiencia en la identificación de características no verbales clave para evaluar la atención en tiempo real.  2. Altos requerimientos computacionales de los modelos de IA existentes que impiden su despliegue en dispositivos Edge..  3. Mecanismo de retroalimentación requeridos en tiempo real y usable que convierta los datos de atención en información accionable para entender el comportamiento y compromiso de atención de los estudiantes. | **Objetivos específicos**  1. Identificar, seleccionar y extraer las características faciales y posturales más representativas que se asocian directamente con los estados de atención, distracción y fatiga en estudiantes universitarios.  2. Implementar y evaluar técnicas de optimización (ej. cuantización) en los modelos ligeros de detección para reducir la carga computacional, el consumo de energía y la latencia, asegurando un rendimiento óptimo en dispositivos Edge.  3. Evaluar el desempeño del prototipo funcional de Edge AI en condiciones reales de aula híbrida, validando su precisión, latencia y la utilidad pedagógica percibida por los docentes (ej. a través de una escala Likert). | **Hipótesis especificas**  1. Las características faciales (ej. mirada, emociones) y posturales (ej. inclinación corporal, posición de cabeza) seleccionadas permiten distinguir con una alta precisión los estados de atención, distracción y fatiga.  2. La optimización de los modelos mediante cuantización y/o poda mantendrá la precisión de detección y logrará una reducción significativa de la carga computacional.  3. El sistema Edge AI mejorará significativamente la capacidad del docente para monitorear y responder a la atención estudiantil en tiempo real, obteniendo un nivel de usabilidad percibida superior a 4/5 en la escala Likert.  . | **Variable dependiente:**  Nivel de atención estudiantil  **Dimensiones:**   * Expresiones faciales (Atento, distraído, fatigado). * Postura corporal (inclinación, mirada, orientación). * Índice de atención.   **Métricas de Evaluación:**  Precisión, Latencia, Frames por Segundo (FPS), Usabilidad (Escala Likert). | **Población y muestra:**  Docentes y estudiantes de aulas híbridas (a definir). |
| **Técnicas e Instrumentos de recolección de datos:** |
| **Técnicas e Instrumentos de análisis y procesamiento de datos:** |