

Reto Modelo y Refinamiento

Juan Pablo Castañeda Serrano
Aldo Daniel Villaseñor Fierro
José Alfredo García Rodríguez
Francisco Castorena Salazar

A01752030
A01637907
A00830952
A00827756

Contexto de la problemática y objetivos

Monitoreo de asistencias y participaciones de alumnos en un salón de clases por medio de herramientas y técnicas de Computer Vision, principalmente con face recognition y pose detection.

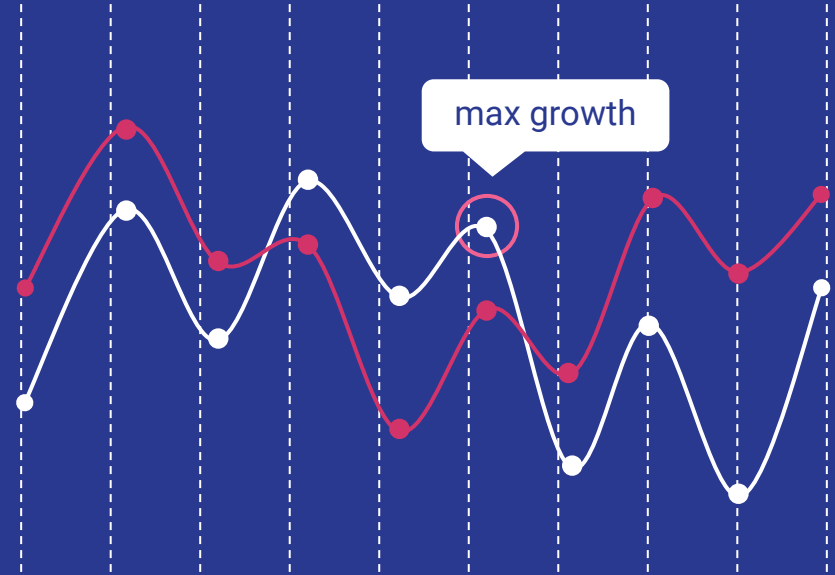
Esto para automatizar y tener un mejor control de asistencia, garantizar una participación activa y tener datos cuantitativos y estadísticas relevantes sobre los alumnos.

Flujo de trabajo



Estado del arte

Alternativas de solución



Yolov8 (You Only Look Once)

- Es un sistema de detección de objetos en imágenes y videos.
- Este ha sido destacado por su velocidad y precisión.
- Utiliza CNN para predecir la ubicación e identificar tipos de objetos simultáneamente, la arquitectura de esta red está hecha de tal forma que permite identificar objetos en diferentes escalas.



Yolov8 (You Only Look Once)

- Yolo fue desarrollado y entrenado utilizando Darknet, un framework escrito en C y CUDA para el desarrollo de redes neuronales.
- Originalmente Yolo se entrenó con el conjunto de datos de COCO (Common Objects in Context) que contiene una amplia variedad de imágenes etiquetadas con múltiples clases de objetos.

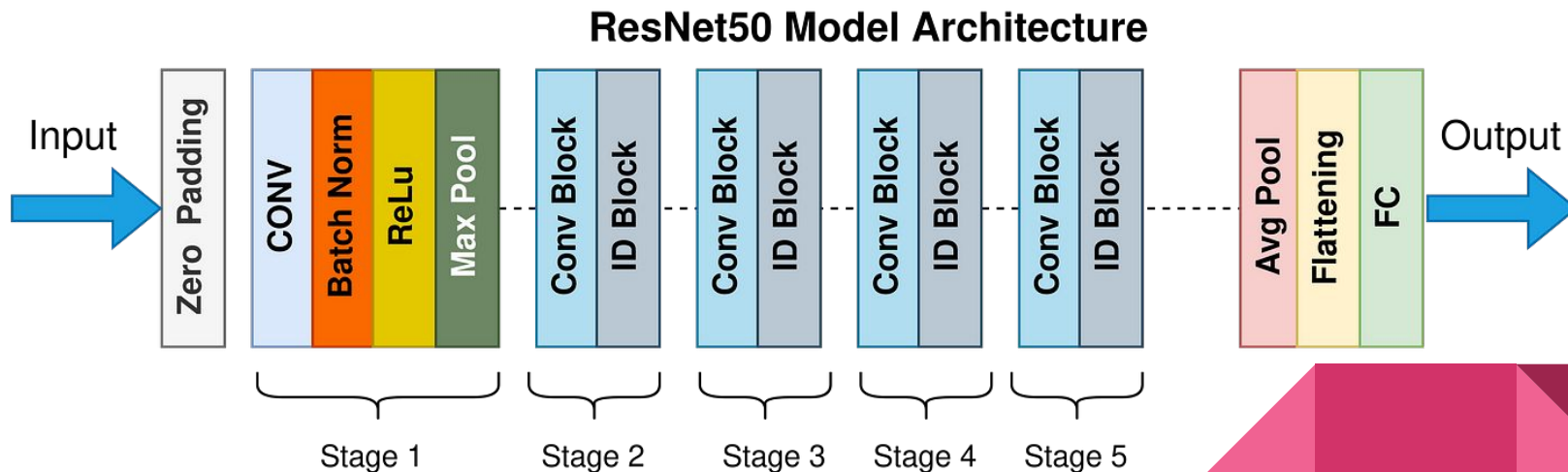


COCO

Common Objects in Context

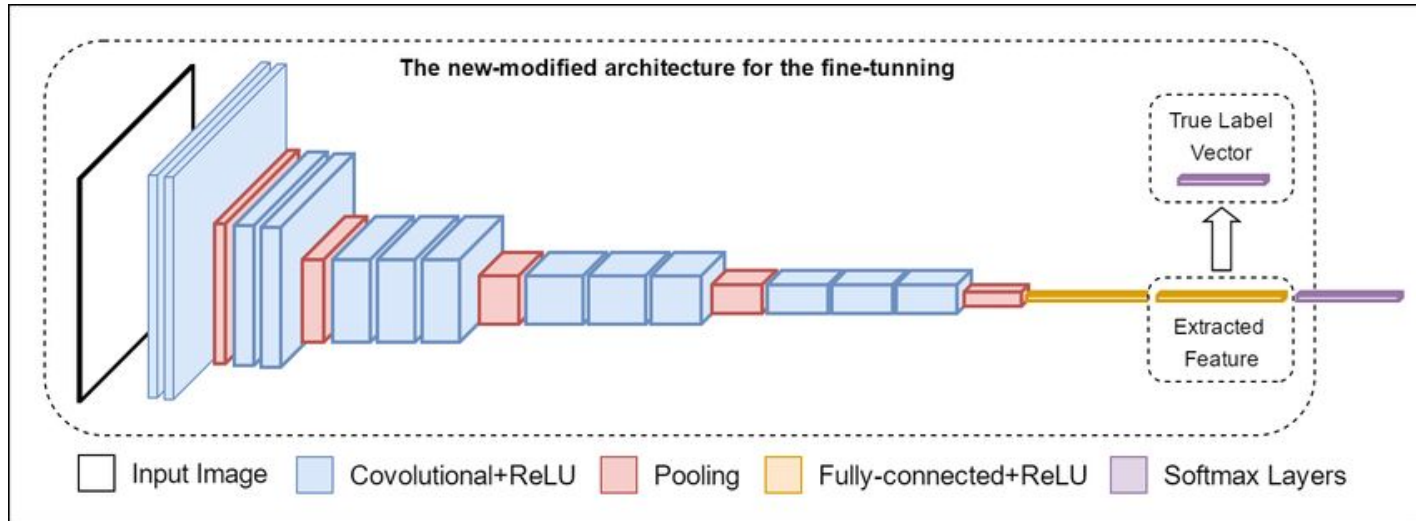
face-recognition

- Esta biblioteca proporciona una interfaz sencilla para detectar rostros en imágenes, reconocer rostros y extraer características faciales. Utiliza otras bibliotecas populares como dlib y OpenCV en su implementación. Dlib usa ResNet como arquitectura.



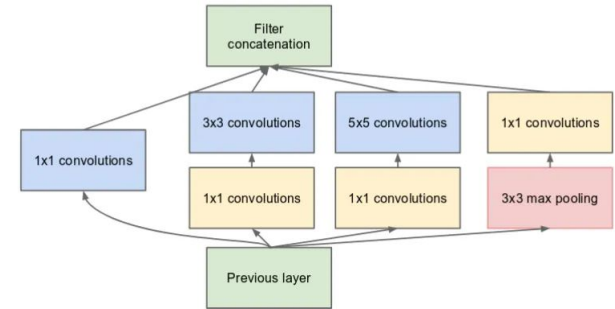
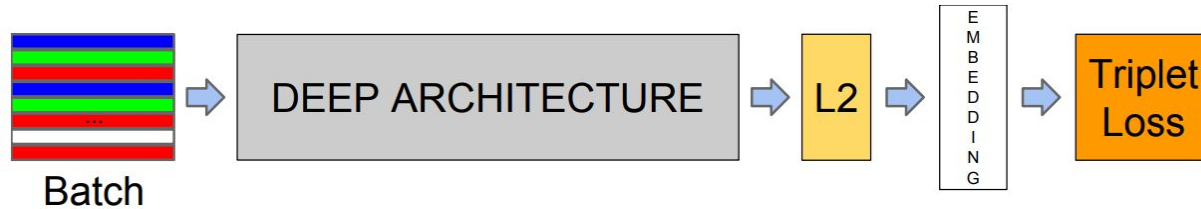
VGG-16 Face

- Es un modelo de reconocimiento facial desarrollado en la Universidad de Oxford. Este modelo se basa en la arquitectura VGGNet, utilizando capas convolucionales y capas completamente conectadas para aprender representaciones discriminatorias de las imágenes.



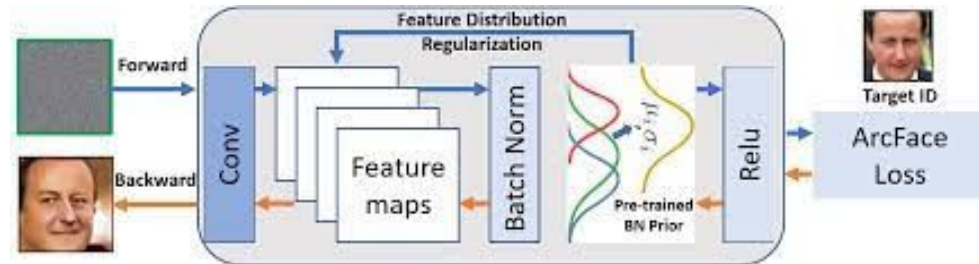
Google - FaceNet

- Es un modelo de reconocimiento facial desarrollado en la Universidad de Oxford. Este modelo se basa en la arquitectura VGGNet, utilizando capas convolucionales y capas completamente conectadas para aprender representaciones discriminatorias de las imágenes.



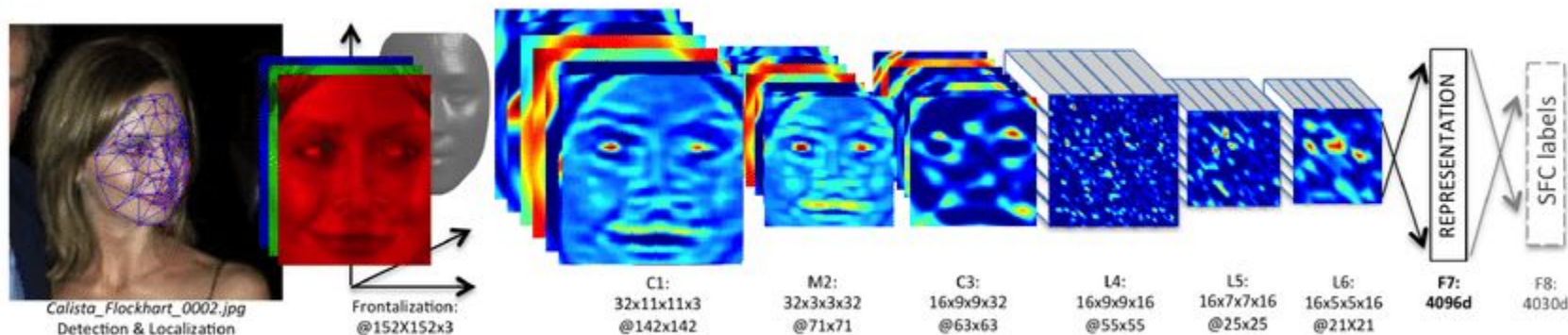
ArcFace

- Se basa en la arquitectura de CNN y es conocido por su capacidad para generar embeddings altamente discriminativos y robustos. Este método fue introducido para mejorar la precisión y la confiabilidad del reconocimiento facial, en situaciones en las que las variaciones de iluminación, expresiones faciales y ángulos de visión son comunes.



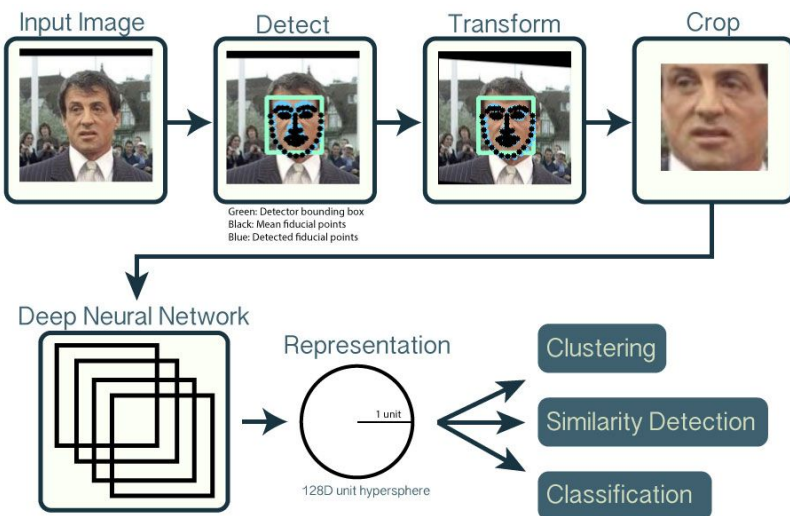
Facebook (Meta) - DeepFace

- Es un sistema de reconocimiento facial desarrollado por Facebook que utiliza una CNN para analizar imágenes faciales. Fue creada para mejorar la precisión del reconocimiento facial en condiciones desafiantes.



OpenFace

- Es una biblioteca de código abierto desarrollada por Carnegie Mellon University que utiliza técnicas de aprendizaje profundo para analizar imágenes faciales y extraer características significativas de las caras. Esta librería ha sido diseñada para ser fácilmente accesible y utilizada por desarrolladores y científicos.



Reconocimiento de Rostros

face-recognition

Esta biblioteca proporciona una interfaz sencilla para detectar rostros en imágenes, reconocer rostros y extraer características faciales. Utiliza otras bibliotecas populares como dlib y OpenCV en su implementación.

VGG16-Face

Es un modelo de reconocimiento facial desarrollado en la Universidad de Oxford. Este modelo se basa en la arquitectura VGGNet, utilizando capas convolucionales y capas completamente conectadas para aprender representaciones discriminatorias de las imágenes.

FaceNet

Es un modelo desarrollado para el reconocimiento facial que se destaca por su capacidad para generar representaciones precisas y compactas de las caras humanas. Fue desarrollado por investigadores de Google en 2015. FaceNet utiliza una arquitectura de red neuronal conocida como siamesa.

Reconocimiento de Rostros

ArcFace

Se basa en la arquitectura de CNN y es conocido por su capacidad para generar embeddings altamente discriminativas y robustas. Este método fue introducido para mejorar la precisión y la confiabilidad del reconocimiento facial, en situaciones en las que las variaciones de iluminación, expresiones faciales y ángulos de visión son comunes.

DeepFace

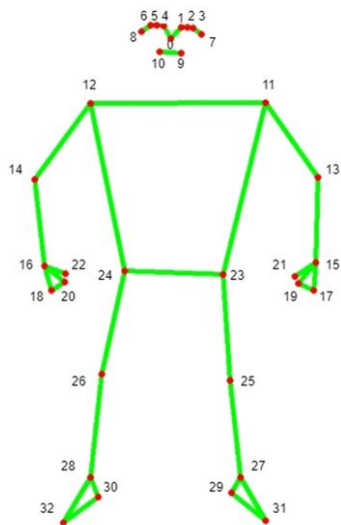
Es un sistema de reconocimiento facial desarrollado por Facebook que utiliza una CNN para analizar imágenes faciales. Fue creada para mejorar la precisión del reconocimiento facial en condiciones desafiantes.

OpenFace

Es una biblioteca de código abierto desarrollada por Carnegie Mellon University que utiliza técnicas de aprendizaje profundo para analizar imágenes faciales y extraer características significativas de las caras. Esta librería ha sido diseñada para ser fácilmente accesible y utilizada por desarrolladores y científicos.

Reconocimiento de Posturas

- El reconocimiento de posturas en nuestro caso será hecho con MediaPipe. Esta utiliza modelos de redes neuronales pre-entrenadas para comprender y predecir la posición de las articulaciones del cuerpo humano.



 **MediaPipe**

Almacenamiento y Acceso a Datos

localhost:27017

My Queries

Databases

Search

- admin
- config
- fcRecog**
 - Poses
 - estudiantes
 - images
- local

Collections

+ Create collection

Refresh

View

Sort by

Collection Name

estudiantes

Storage size:

331.78 kB

Documents:

4

Avg. document size:

76.61 kB

Indexes:

1

Total index size:

36.86 kB

images

Storage size:

1.68 MB

Documents:

38

Avg. document size:

43.80 kB

Indexes:

1

Total index size:

36.86 kB

Poses

Storage size:

4.10 kB

Documents:

0

Avg. document size:

0 B

Indexes:

1

Total index size:

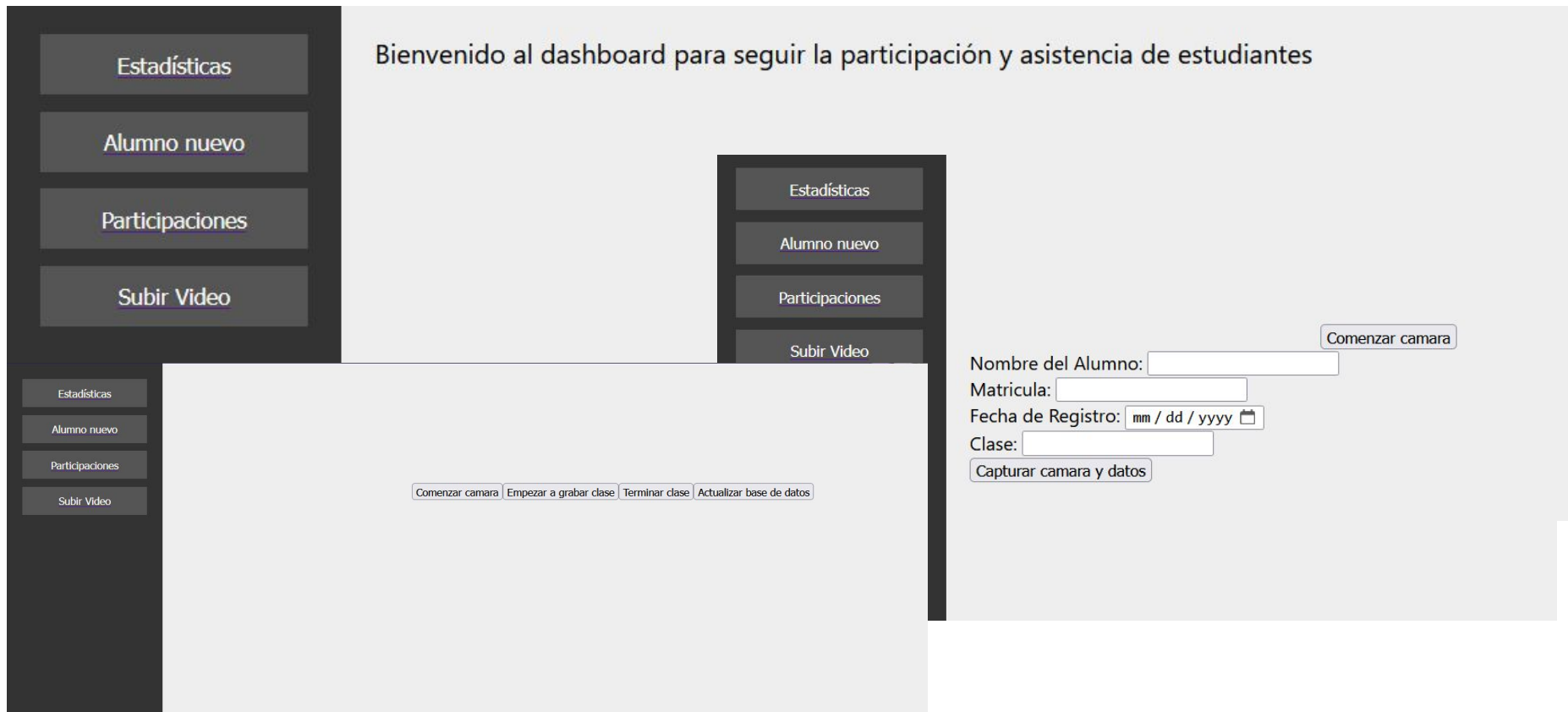
4.10 kB

Estudiantes

Imágenes

Participaciones

Interfaz Web



Transfer Learning YOLOv8 nano

Entrenamiento de la Arquitectura de YOLO

Box_loss=0.9382

Reentrenamiento de YOLO 10 épocas

Reentrenamiento de YOLO 20 épocas