

INGELIGENCIA ARTIFICIAL II

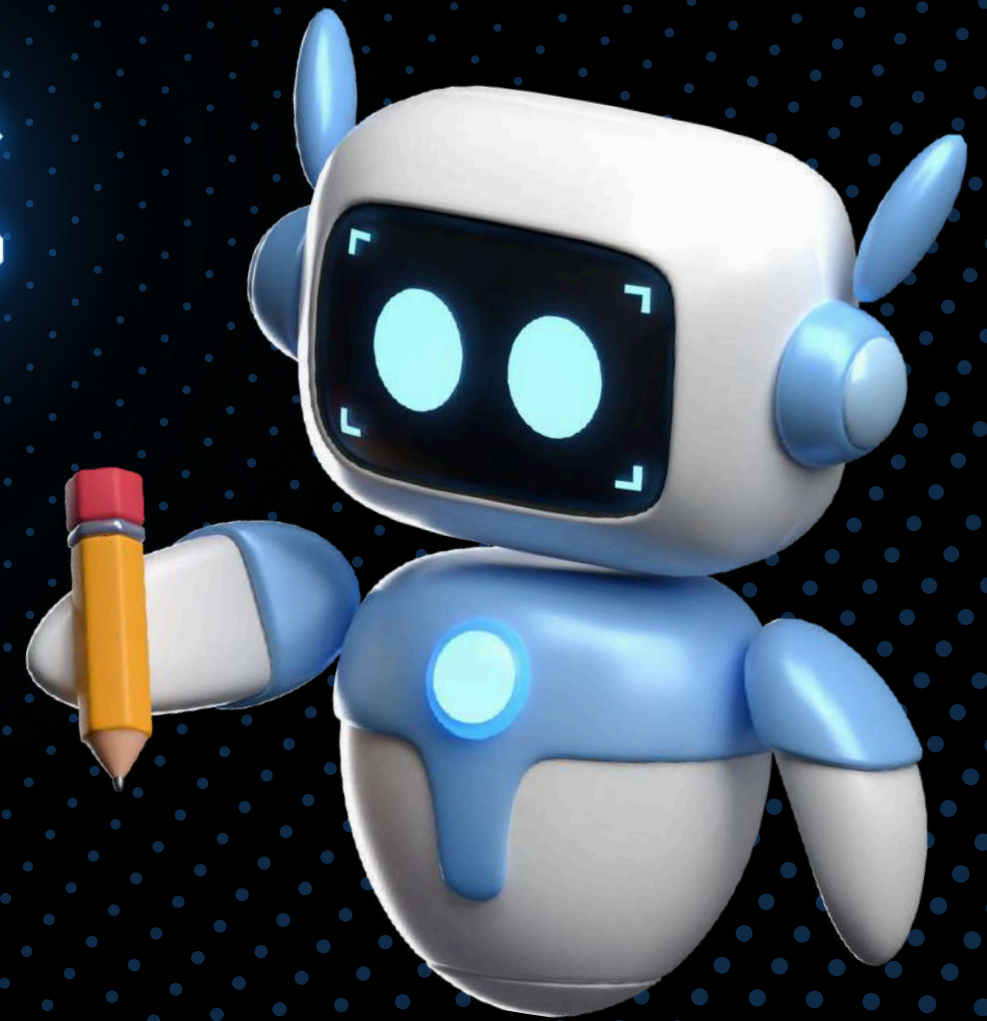
FacePass



JOHAN SEBASTIÁN
LEÓN PEÑALOZA



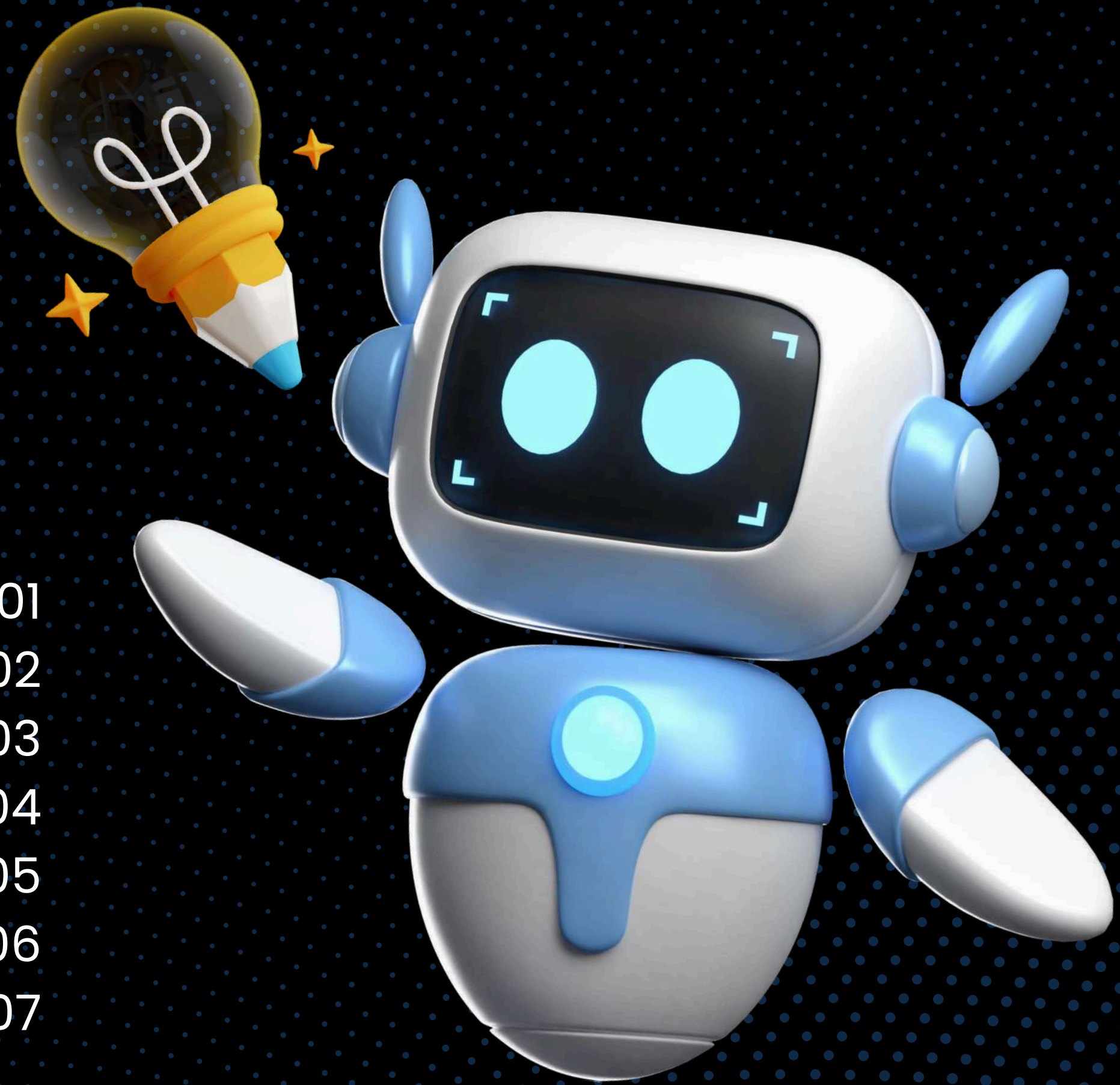
CRISTIAN STIVENS VILLAREAL PARRA



JOAN SEBASTIÁN
PATIÑO JAÍMES

TABLA DE CONTENIDO

• Introducción	01
• Objetivos del proyecto	02
• Alcance del proyecto	03
• Datasets	04
• Metodología	05
• Resultados preliminares	06
• Conclusión	07



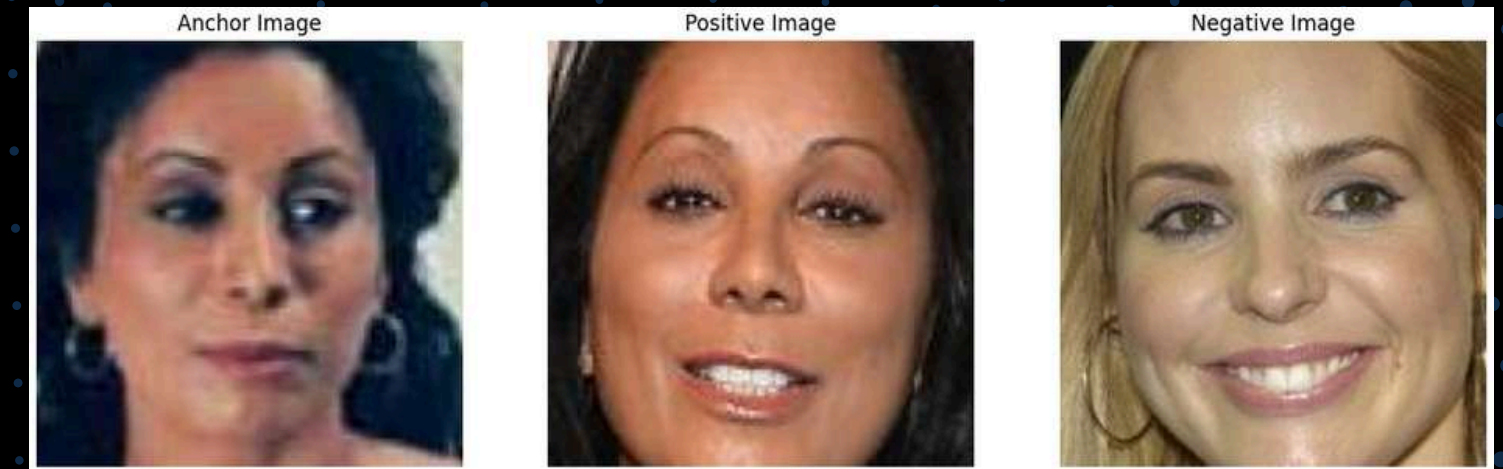
INTRODUCCIÓN



Al intentar ingresar al estadio para ver un partido del Bucaramanga, el sistema indicó que alguien con el mismo número de cédula ya había ingresado. Esto generó sospechas de clonación de mi cédula, resultando en una acusación de falsificación y en la ocupación de mi asiento por otra persona, causando una situación incómoda y frustrante.

¿Validación de identidad de forma tradicional?

Este incidente evidencia un problema recurrente en el uso de métodos tradicionales como la validación por medio de documentos físicos (cédulas), los cuales son vulnerables a la clonación y fraudes, especialmente en eventos masivos.



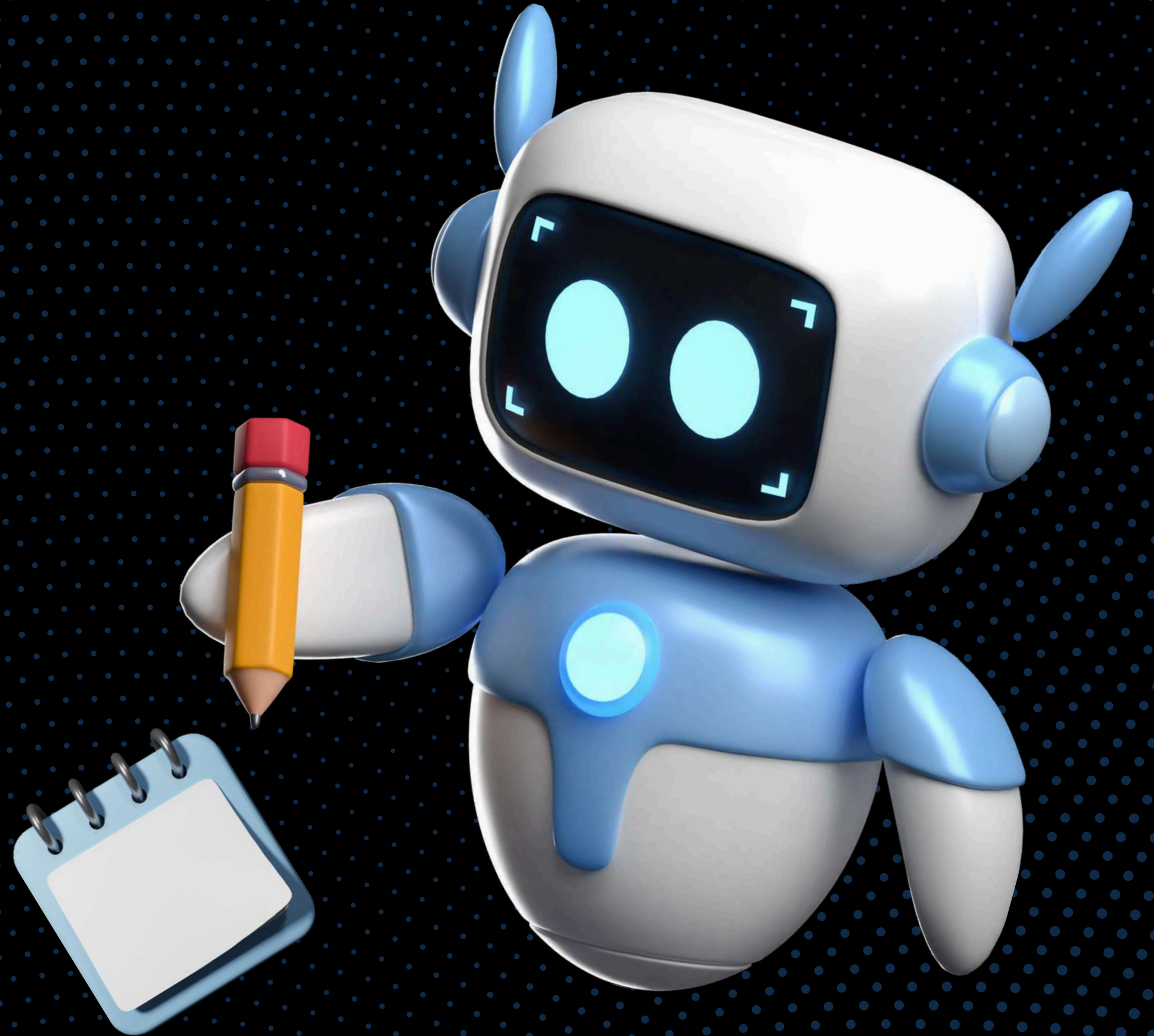
OBJETIVO DEL PROYECTO

Implementar un modelo de validación de rostros mediante redes convolucionales para la correcta gestión de ingreso de las personas en distintos escenarios.



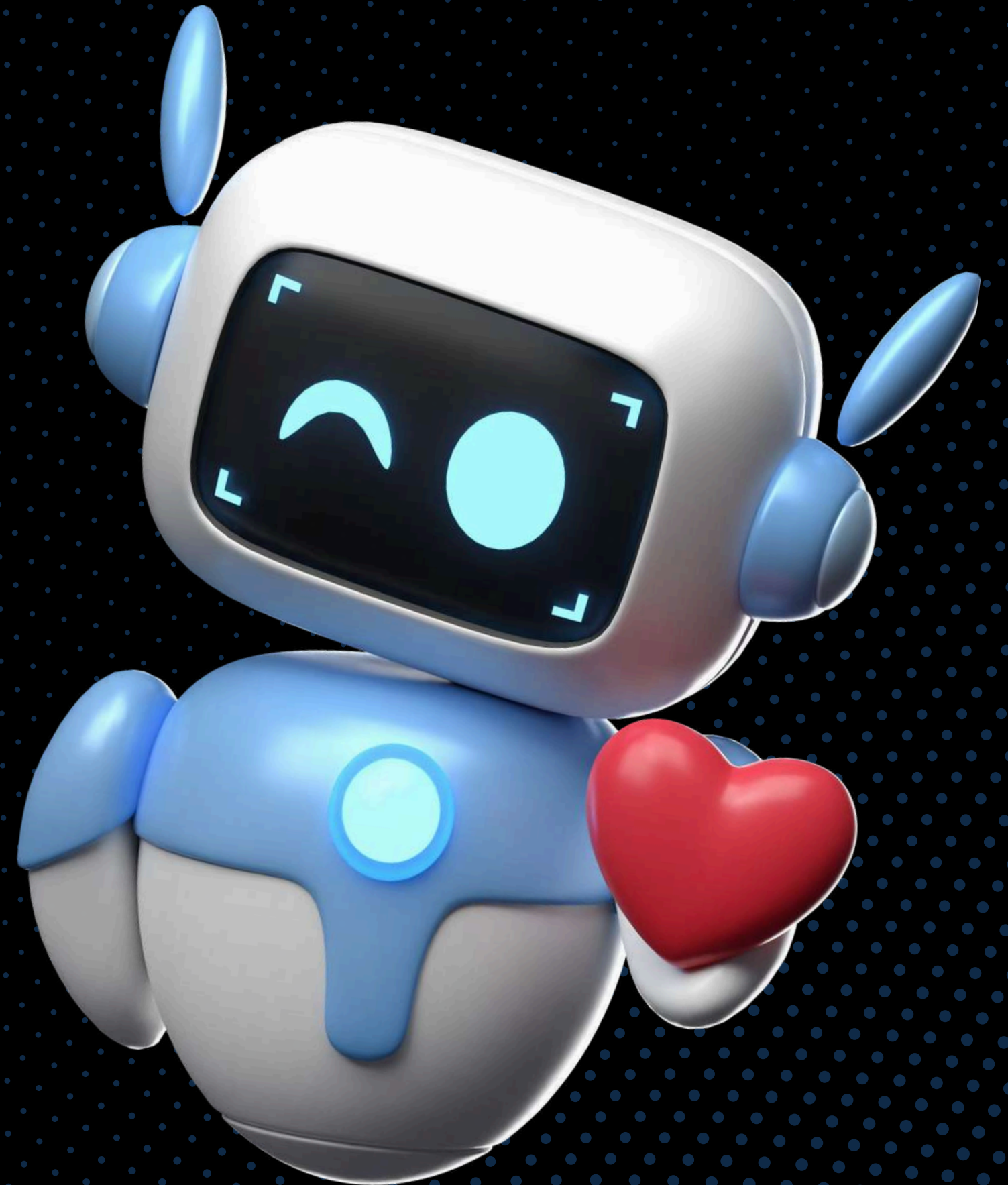
ALCANCE DEL PROYECTO

Que el modelo sea capaz de diferenciar correctamente rostros que cumplan con los parámetros de validación.



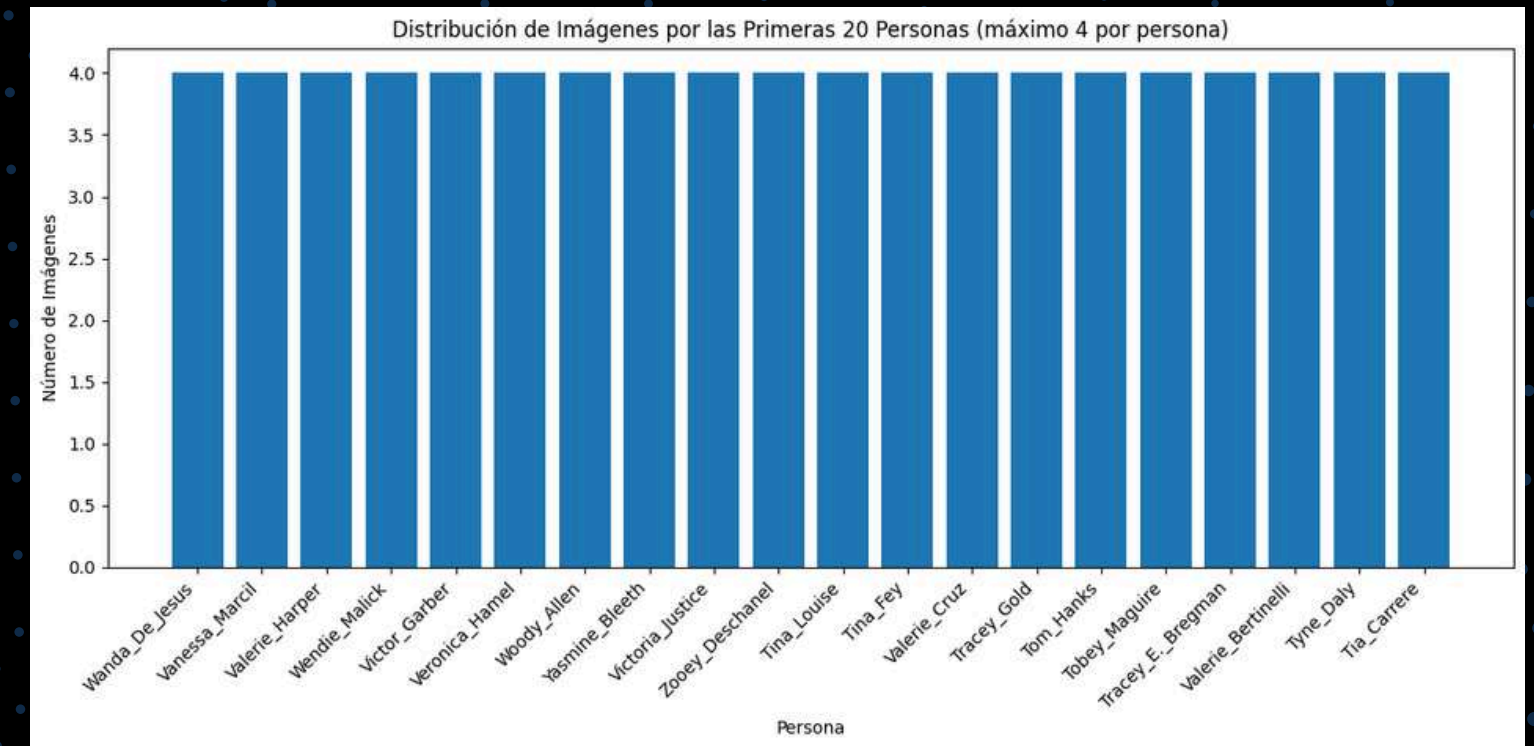
DATASET

FACESCRUB



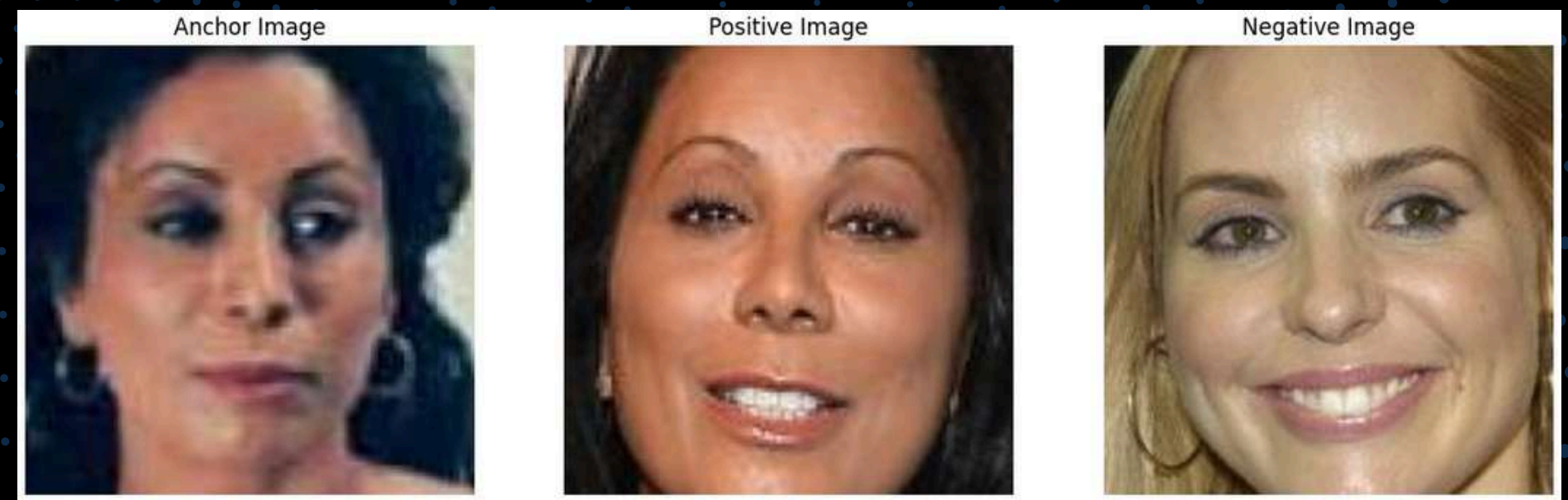
PRE-PROCESAMIENTO SIAMES

- Primeras 4 imágenes por carpeta (límite total de 4000).
- Imágenes leídas con OpenCV; descarta imágenes fallidas.
- Escala de Grises: se convirtieron las imagenes para simplificar características.
- Tamaño fijo: 100x100 píxeles.
- Normalización: Escala de valores: 0 a 1 (división por 255).
- Almacenamiento Final: Conversión a arrays de NumPy para facilitar el entrenamiento.



PRE-PROCESAMIENTO SIAMÉS CON TRIPLET LOSS

- Primeras 4 imágenes por carpeta (límite total de 4000).
- Imágenes leídas con OpenCV; descarta imágenes fallidas.
- Tamaño fijo: 224x224 píxeles.
- Normalización: Escala de valores: 0 a 1 (división por 255).
- Almacenamiento Final: Conversión a arrays de NumPy para facilitar el entrenamiento.



METODOLOGÍA

¿Siamese Network?

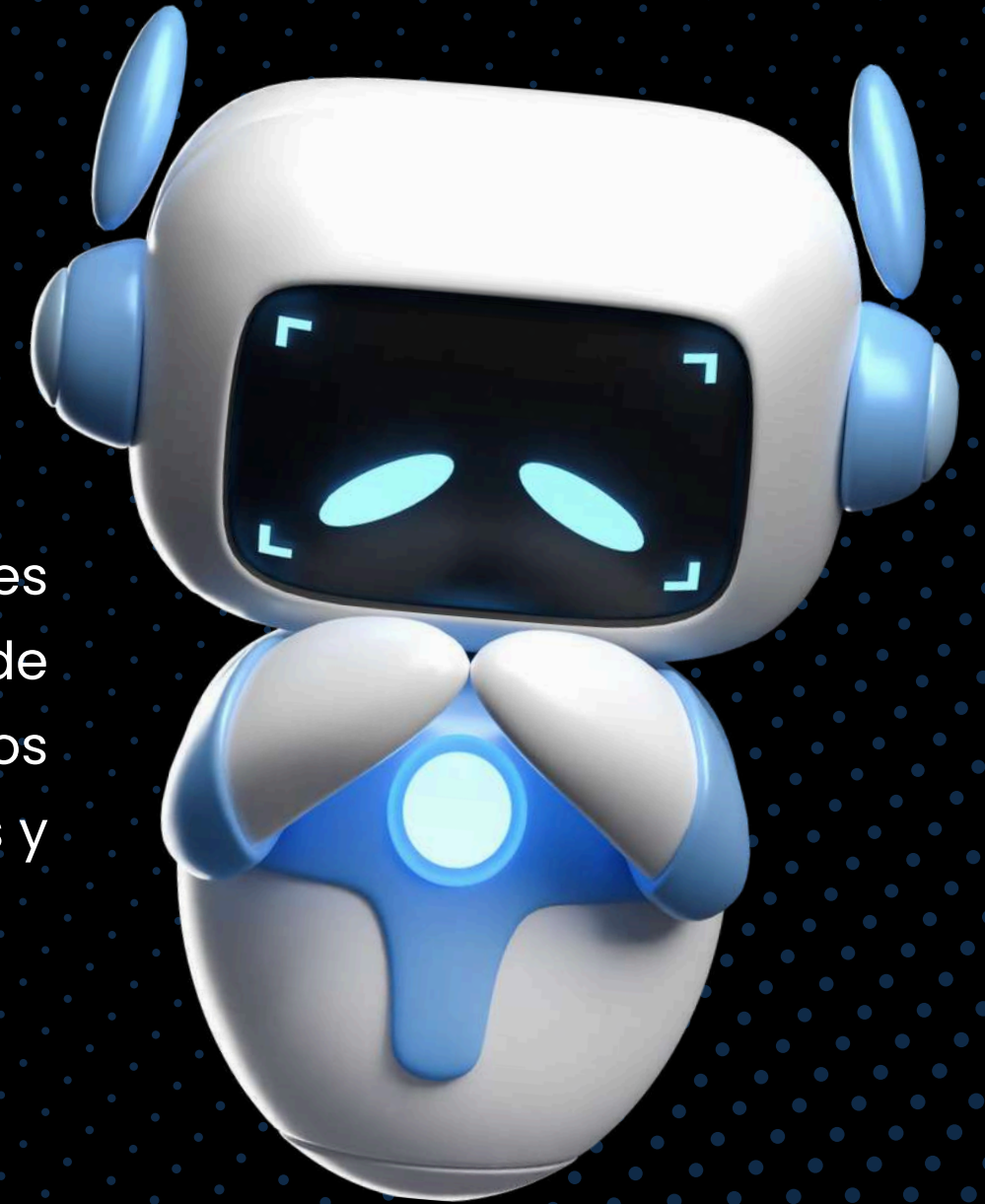
Red neuronal que se utiliza para **comparar dos entradas** y determinar cuán similares son entre sí. La red toma dos imágenes como entrada (en este caso dos imágenes de rostros) y pasa cada una de ellas a través de una subred idéntica, compartiendo los mismos parámetros. Después, la red produce una representación de las dos imágenes y calcula la “similitud” mediante alguna función de distancia (Euclidiana o L1).

Contrastive loss

$$L = \frac{1}{2N} \left(\sum_{n=1}^N y_n d_n^2 + (1 - y_n) \max(\text{margin} - d_n, 0)^2 \right)$$

where $d_n = \|\mathbf{x}_{0n} - \mathbf{x}_{1n}\|_2$, N denotes the mini-batch size.

(from <http://docs.chainer.org/en/stable/reference/functions.html#chainer.functions.contrastive>)



METODOLOGÍA

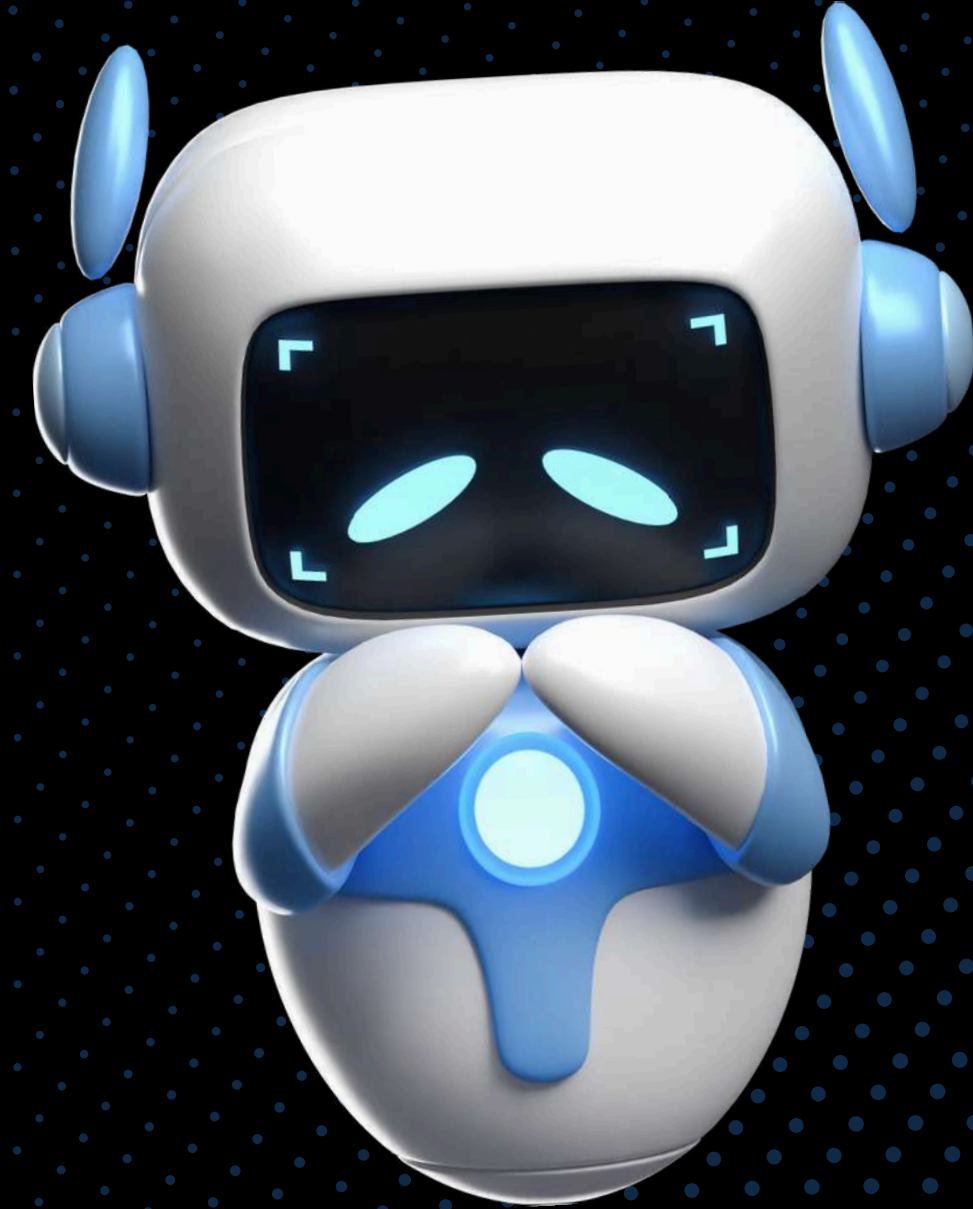
Triplet Loss

La triplet loss es una técnica utilizada en redes neuronales, especialmente en tareas de verificación e identificación de imágenes, como en el caso del reconocimiento facial. El objetivo de este enfoque es entrenar un modelo para aprender representaciones que puedan comparar imágenes entre sí y determinar si son similares o diferentes.

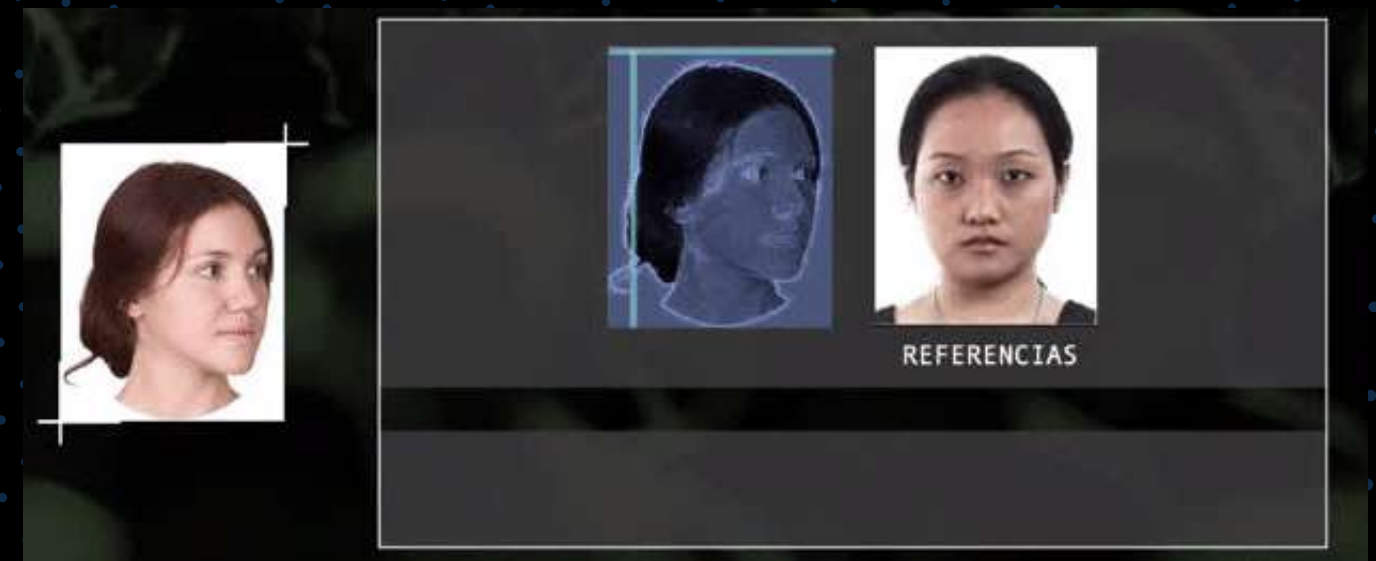
¿Qué es un Triplete?

Un triplete es un conjunto de tres imágenes:

- Ancla (Anchor): La imagen de referencia.
- Positiva (Positive): Una imagen de la misma clase que la imagen ancla (por ejemplo, otra foto de la misma persona).
- Negativa (Negative): Una imagen de una clase diferente a la imagen ancla (por ejemplo, una foto de otra persona).



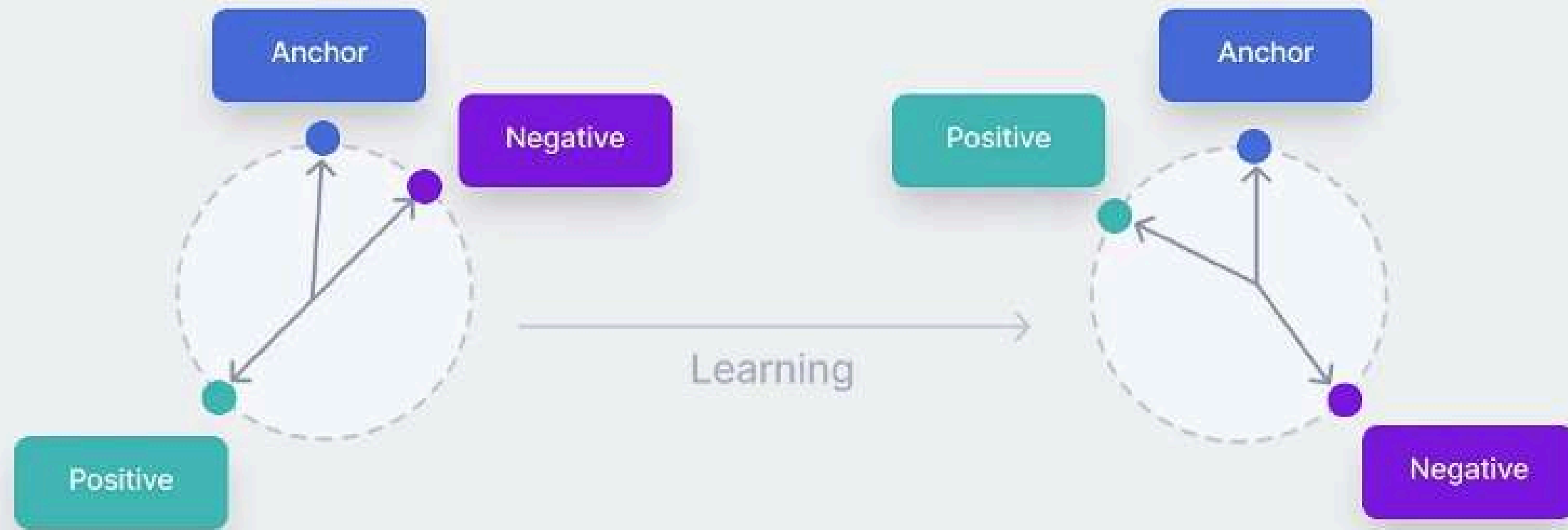
FUNCIÓN DE PÉRDIDA: TRIPLLET LOSS



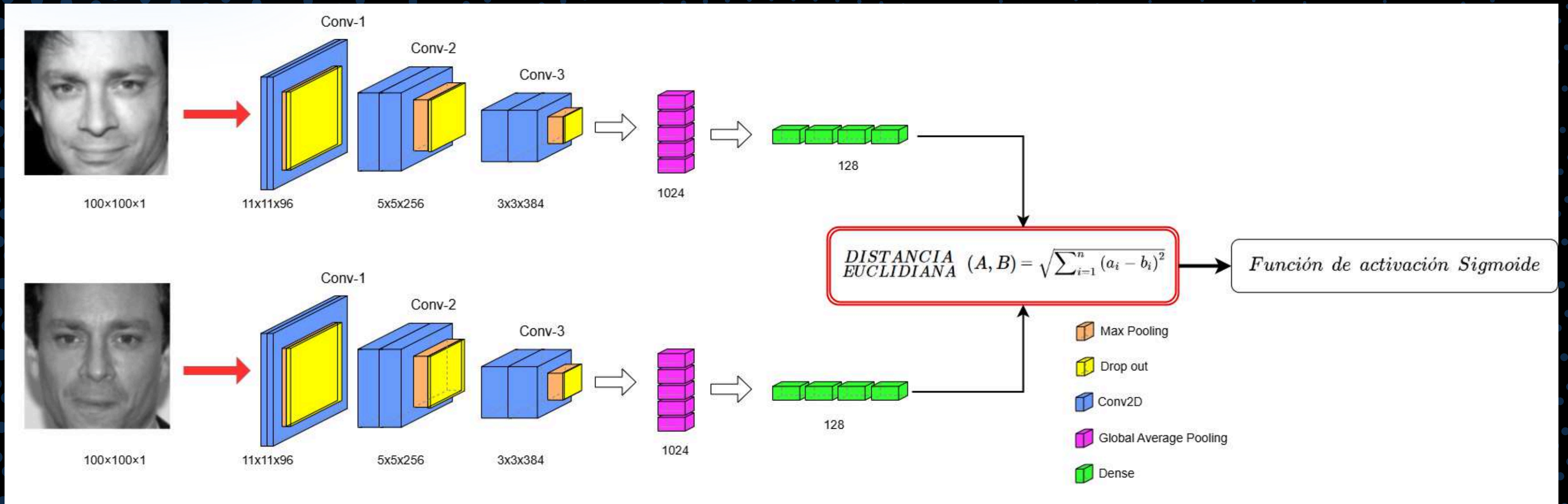
La triplet loss es la función de pérdida utilizada para entrenar el modelo con las tripletas. Su objetivo es asegurarse de que el modelo aprenda representaciones donde las imágenes de la misma clase (ancla y positiva) estén más cerca en el espacio de características y las imágenes de clases diferentes (ancla y negativa) estén más alejadas.

$$L(A,P,N)=\text{MAX}(D(A,P)-D(A,N)+\text{ALPHA},0)$$

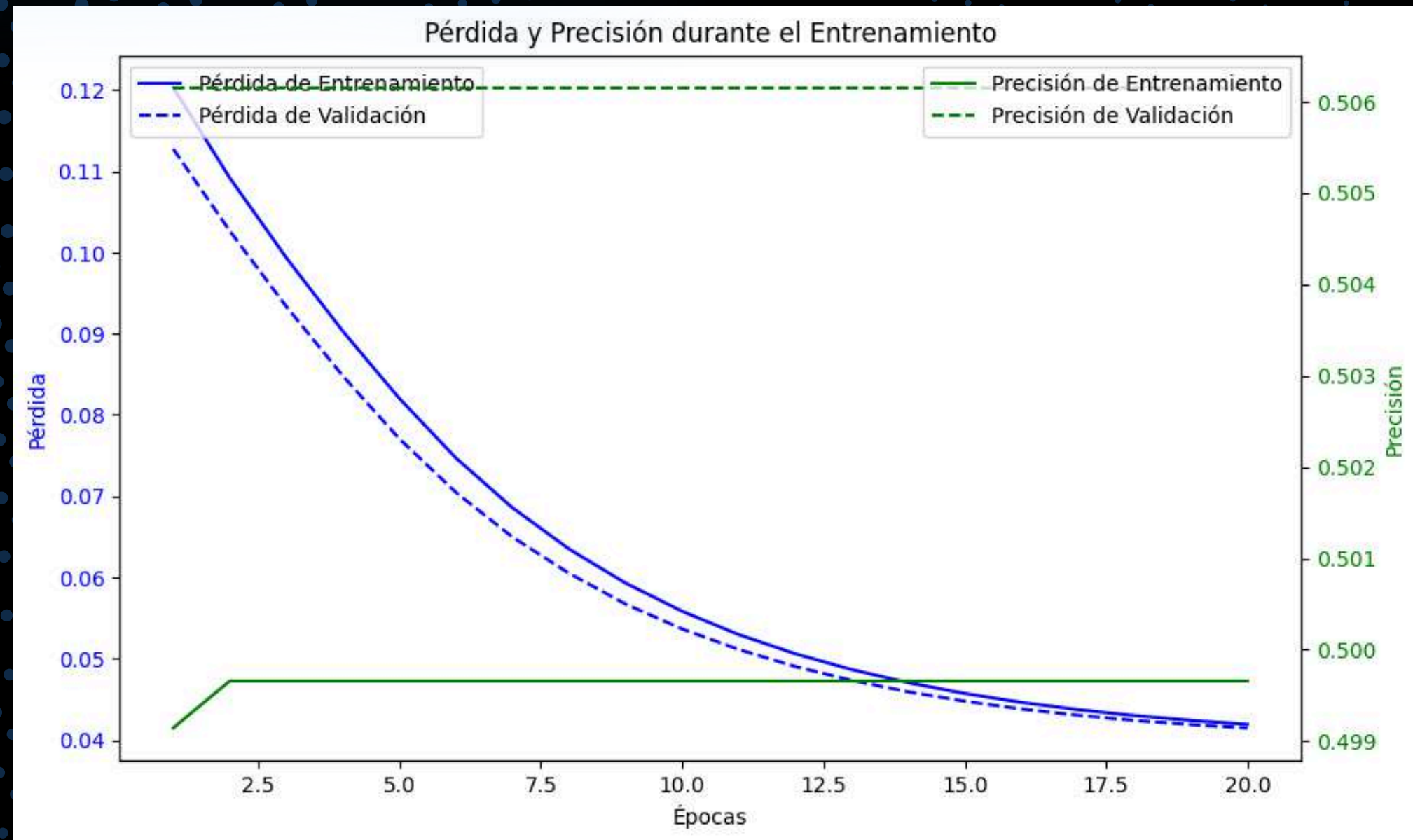
TRIPLER LOSS



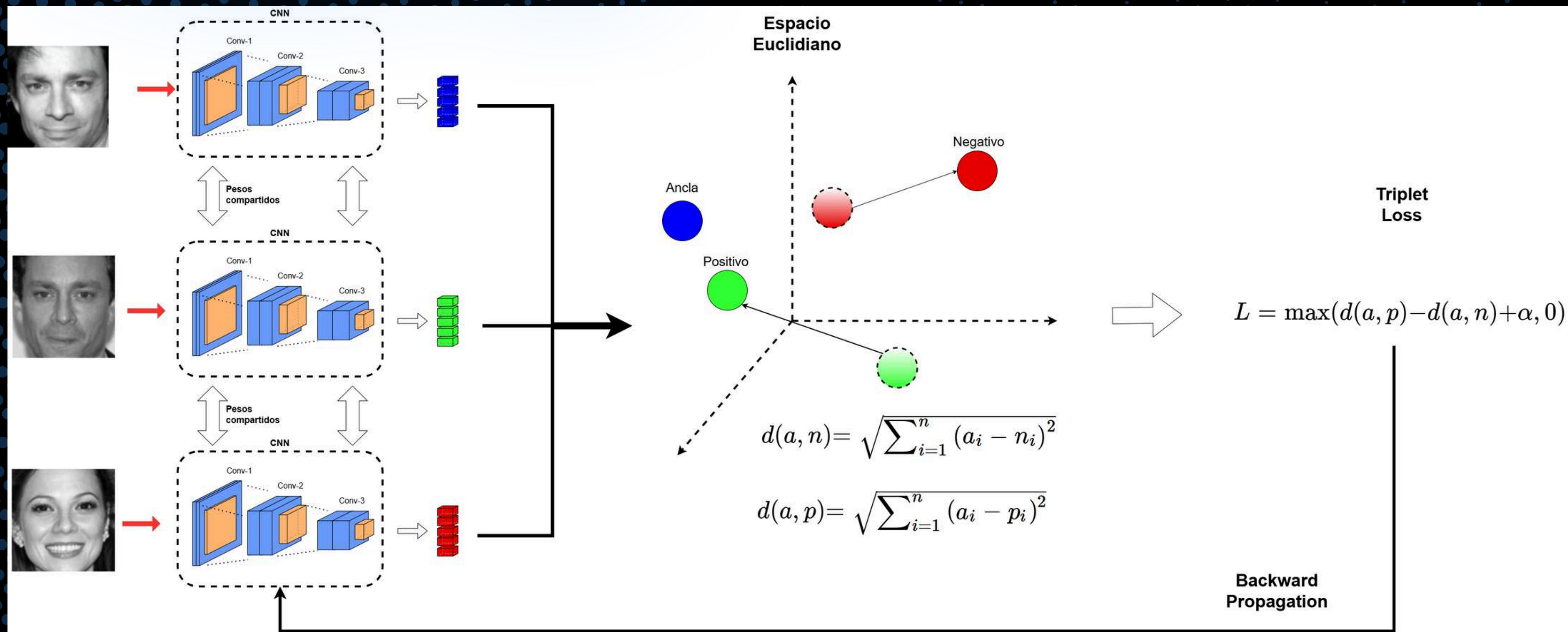
METODOLOGÍA SIAMES



CONTRASTIVE LOSS



METODOLOGÍA SIAMES Y TRIPLET LOSS



DISTANCIA EUCLIDIANA

Es una medida de la distancia entre dos puntos en un espacio vectorial. Dado que los embebidos son vectores en un espacio de alta dimensionalidad, al calcular su distancia Euclidiana entre ellos proporciona una medida de similitud.

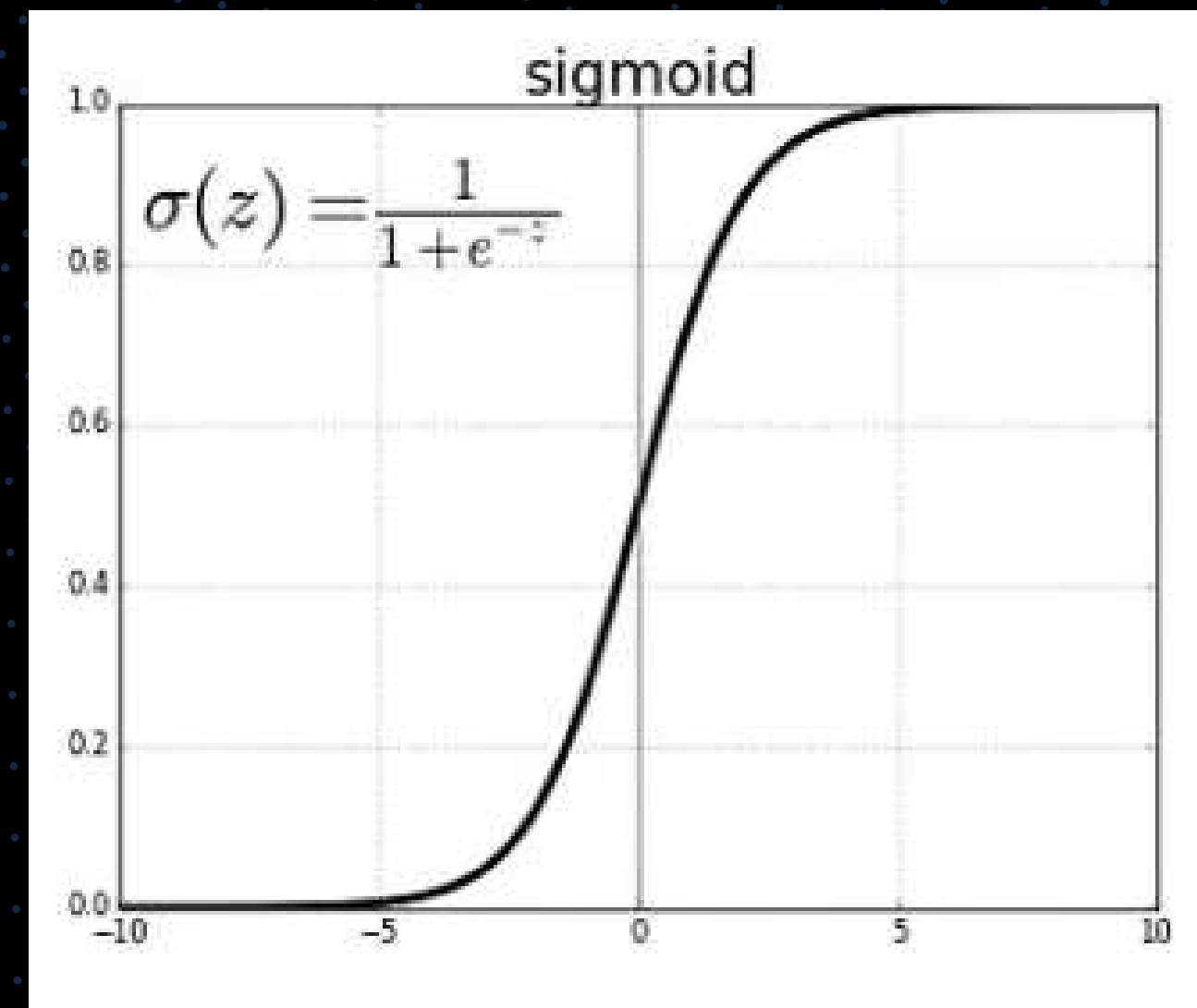
Las redes siamesas son arquitecturas que en lugar de hacer una clasificación tradicional, estas redes generan embebido con representaciones de la imagen. La distancia euclidiana se encarga de comparar dos entradas y determinar su similitud.

$$\textit{DISTANCIA EUCLIDIANA} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

SIGMOIDE

la sigmoide nos ayuda a mapear la distancia entre las representaciones de las dos imágenes hallada en la distancia euclidiana, mapeándolo en un valor entre 0 y 1.

proporciona una probabilidad entre 0 y 1 según un umbral preestablecido (0.5), lo cual es adecuado para tareas de clasificación binaria



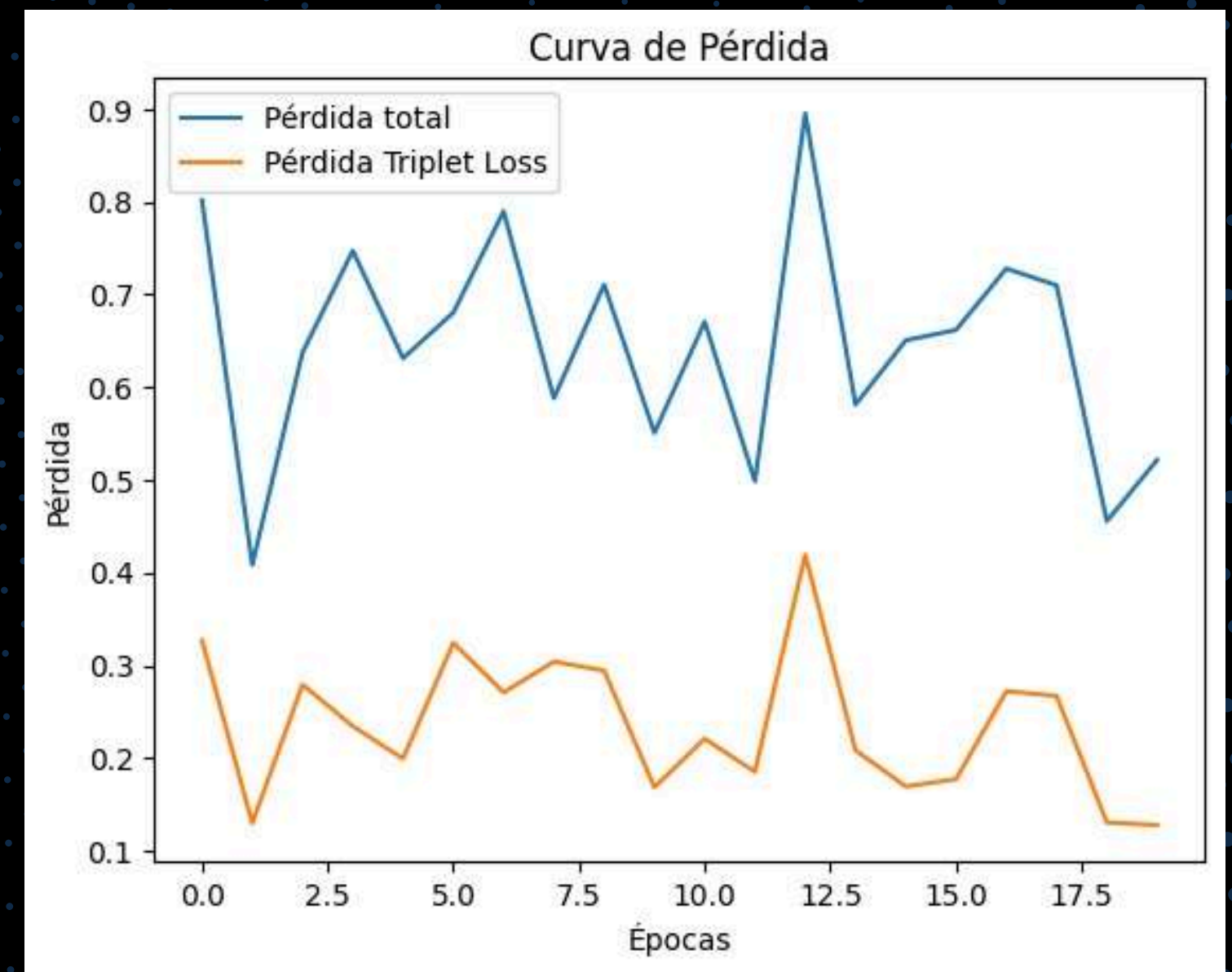
RESULTADOS FACE- NET CON TRIPLETAS

Se obtuvo una variación significativa entre las pérdidas total y Triplet-Loss. Debido a la significativa reducción de los datos de entrenamiento y validación, producto de las limitaciones presentes en colab.

Entrenamiento:

epochs=20,

batch_size=64



CONCLUSIONES

- El primer enfoque de la arquitectura siamese es un buen primer acercamiento hacia cómo funcionan las redes neuronales métricas, sin embargo, la complejidad del dataset puede afectar drásticamente su rendimiento, los problemas de clasificación en validación de rostros con la siamese, enuncian resultados erróneos para el accuracy, sin embargo muchos reportan que su modelo es sólido.
- Probar arquitecturas más complejas y añadir capas de regularización (Dropout/Batch Normalization)
-



"Los mejores emprendedores tienen dos cosas en común: nunca se rinden y nunca paran de aprender".

-John Freddy Vega, CEO de Platzi.

