

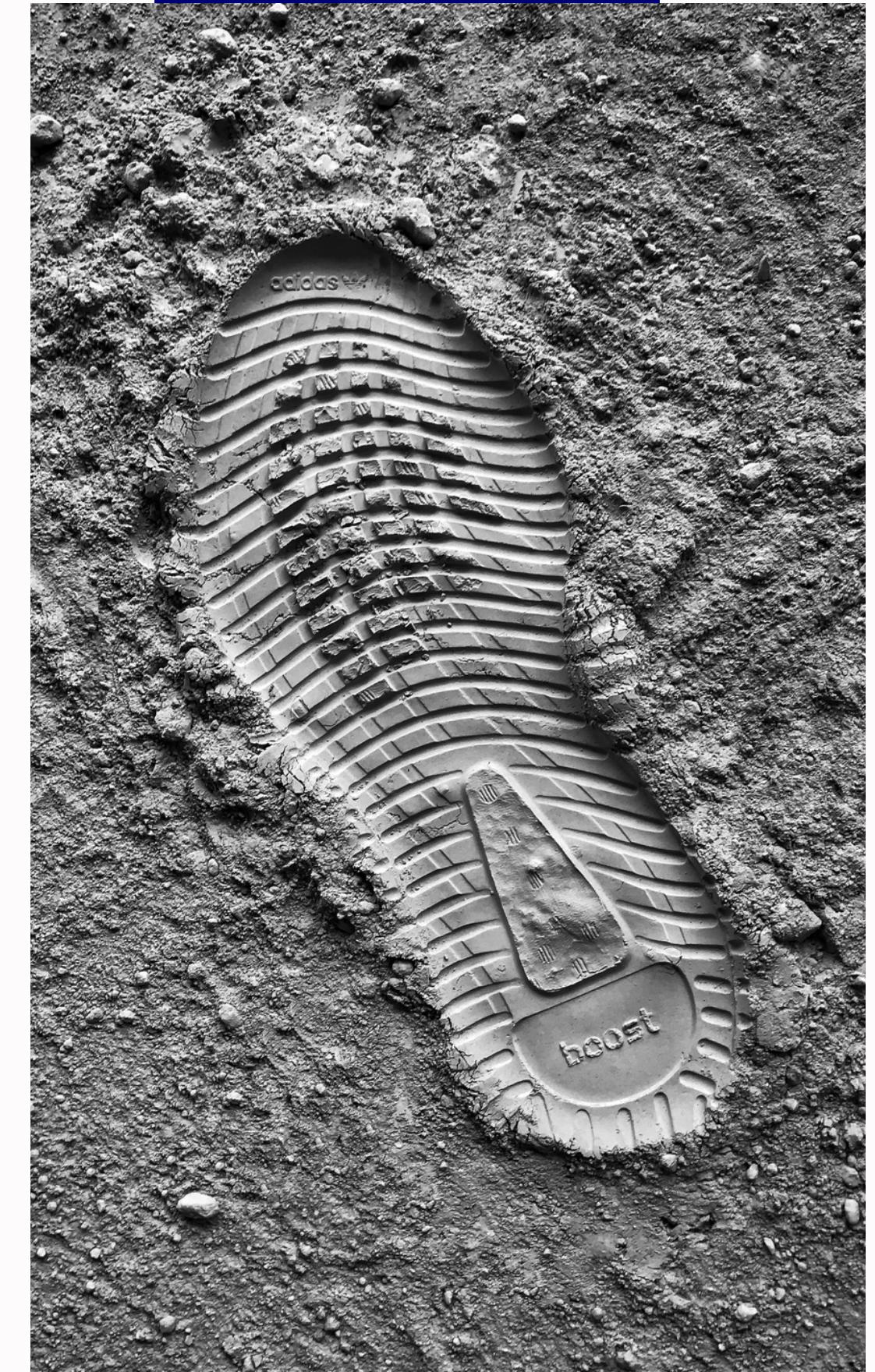
# Identificación de huellas de calzado a partir de imágenes con redes neuronales convolucionales

Máster universitario en Ingeniería computacional y matemática  
Área de Inteligencia Artificial

**Laura Rivera Sánchez**

Tutor/a de TFM  
Elena Álvarez de la Campa Crespo

Profesor/a responsable de la asignatura  
Carles Ventura Royo



# Contexto

*En 2019, INTERPOL efectuó más de 1 600 identificaciones gracias al mayor intercambio y cotejo de huellas dactilares por parte de los países miembros.*

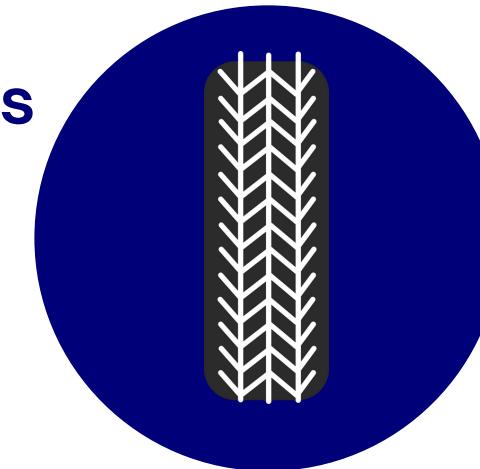
## Huella dactilar

Única para cada persona



## Huella Neumáticos

Cada marca/modelo tiene su relieve



## Huella de zapatos

Bases de datos de huellas en escenas del crimen



## Podología Forense

## Estudio de la marcha



Conocimientos sobre biomecánica para identificar forma de caminar

# Toma de huellas de calzado



*Mediante fotografía  
perpendicular directa  
Uso de guia en "L"*



*Levantamiento de huella  
aplicando yeso*



*Escáner de mano 3d  
captura la topografía*

# Herramientas existentes



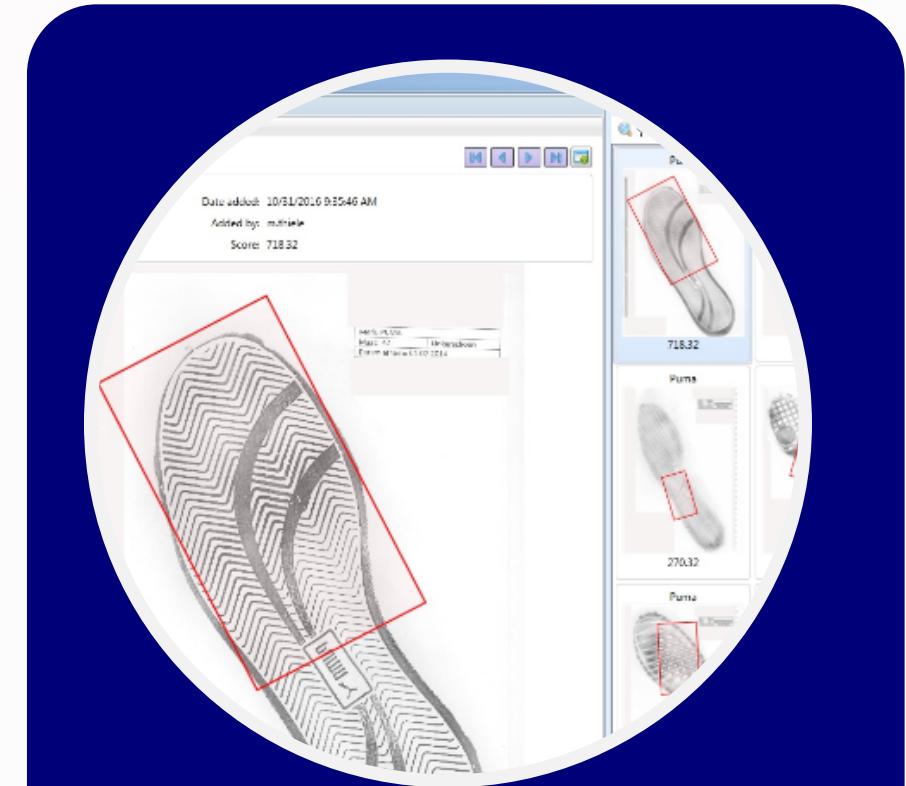
SICAR6

SoleMate 42.000  
muestras.  
Determina marca y  
modelo



DigTrace

Análisis 3D de  
huellas de  
calzado y de  
vertebrados



PRIDE

Comparador  
automático de  
huellas de  
zapatos

# OBJETIVOS



## 1 - Reconocimiento de marca

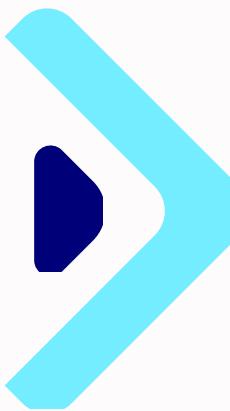
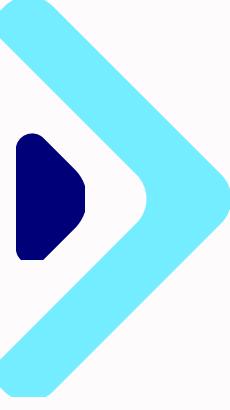
*Red neuronal convolucional de predicción de la marca*



## 2 - Búsqueda de similares

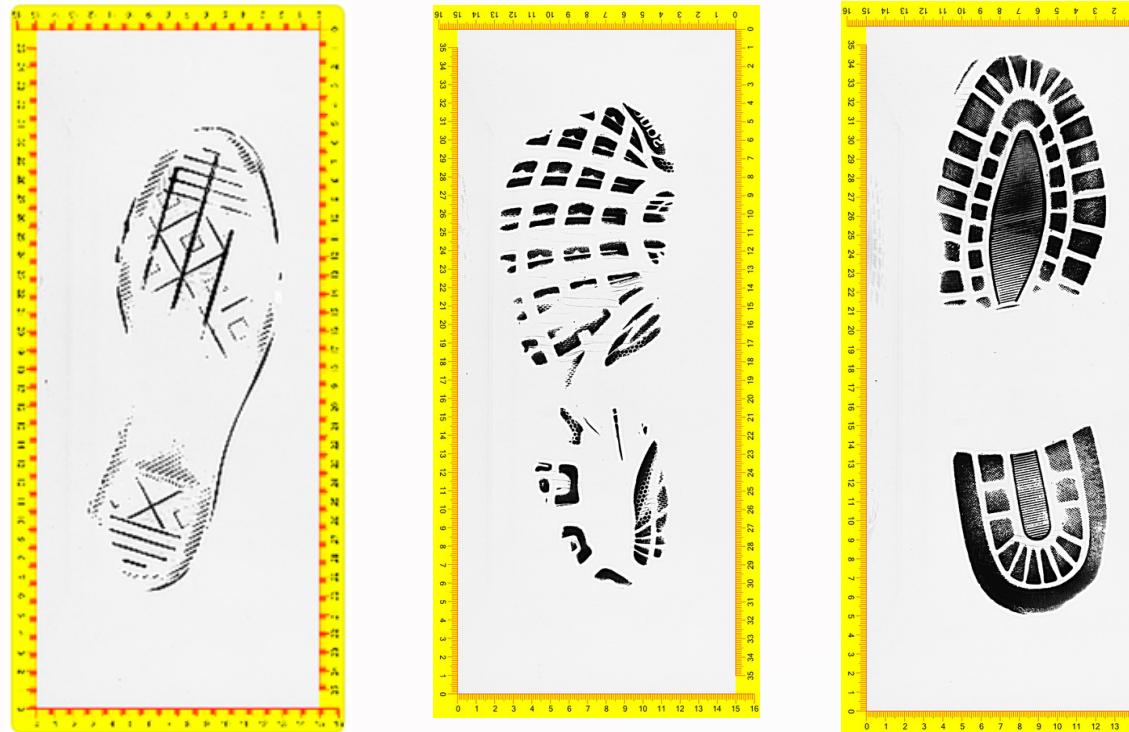
*Red neuronal siamesa para búsqueda de imágenes similares*

# TAREAS PRINCIPALES

- 01  Base de datos de huellas de calzado
- 02  Transformación y aumento de los datos
- 03  Implementación red neuronal convolucional para la predicción de marca
- 04  Experimentos con la red neuronal: transformaciones, iteraciones, muestras, capas y uso de modelo preentrenado
- 05  Implementación red neuronal siamesa para búsqueda de imágenes similares

# Bases de datos

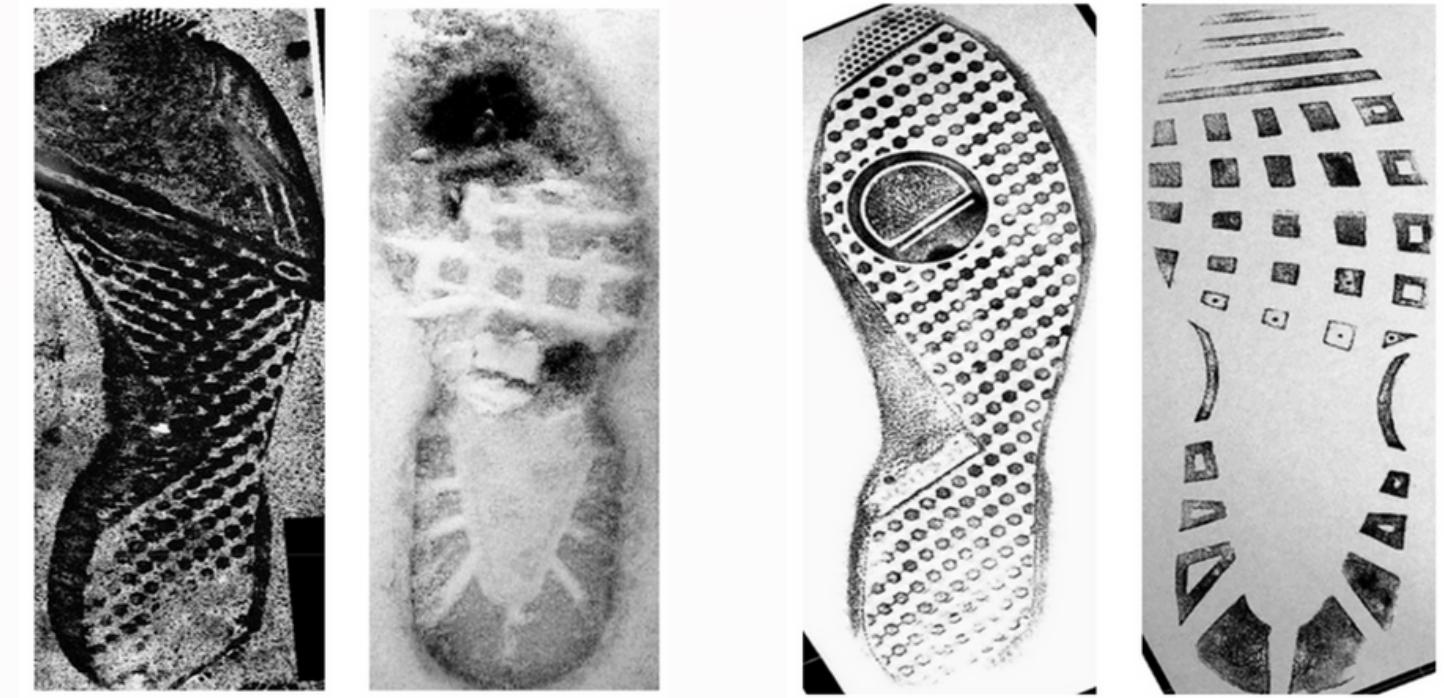
## 2d Footwear



Objetivo 1: Predicción marca

- Universidad Iowa
- 1500 impresiones de 150 muestras
- Datos de genero, marca, modelo, talla

## FD-300



Objetivo 2: Buscar similares

- 300 muestras escenas crimen (izquierda)
- 1175 muestras de referencia (derecha)
- Datos de correspondencia

# Tecnología

01

Lenguaje: Python

02

Jupyter Notebook

03

Objetivo 1: 4 cuadernos

04

Objetivo 2: 1 cuaderno

05

Repositorio: GitHub

06

Macbook Pro Intel i7 Ram 16GB

07

Presentación: Canva



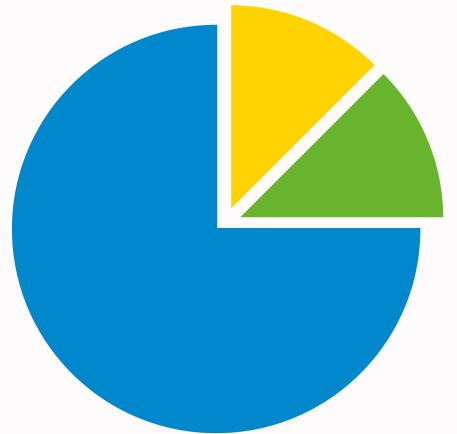
# OBJETIVO 1

Red neuronal  
convolucional para la  
predicción de la marca

# HIPÓTESIS

- *¿Qué preprocessado de imágenes funciona mejor?*
- *¿Utilizar más epochs, resulta siempre en mejor resultado?*
- *¿Un mayor tamaño de las imágenes utilizadas, resulta en un mejor resultado?*
- *¿Cuál es el mínimo de muestras por marca aceptable para el modelo?*

# Análisis de datos 2d Footwear



1500 imágenes

150 registros

28 individuos

62 marcas diferentes

modelo normalmente en blanco

70 mujer - 80 hombre



## Limpieza:

- Marca = "None"
- Marcas con menos de cinco muestras.

\*Muchas marcas con una muestra

# Modelo propuesto

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D) <b>relu</b>	(None, 280, 832, 16)	160
max_pooling2d (MaxPooling2D)	(None, 140, 416, 16)	0
conv2d_1 (Conv2D)	(None, 140, 416, 32)	4640
max_pooling2d_1 (MaxPooling2D)	(None, 70, 208, 32)	0
conv2d_2 (Conv2D)	(None, 70, 208, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, 35, 104, 64)	0
flatten (Flatten)	(None, 232960)	0
dense (Dense)	(None, 128)	29819008
dropout (Dropout)	(None, 128)	0
dense_1 (Dense) <b>softmax</b>	(None, 7)	903
flatten_1 (Flatten)	(None, 7)	0
<hr/>		
Total params:	29,843,207	
Trainable params:	29,843,207	
Non-trainable params:	0	

**Dense:** totalmente conectada

**Conv2D:** flexibilidad y menor tamaño

**MaxPooling2D:** Reduce tamaño y tiempo

**Flatten:** Convierte a una dimensión

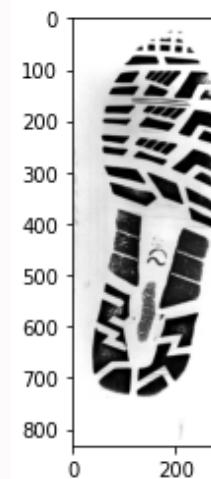
**Dropout:** Se queda con las características más relevantes

Optimizador utilizado: **ADAM**

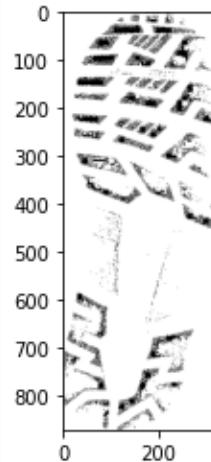
\*Identificación de artistas en cuadros de pintura de N.Viswanathan de la Universidad de Standford  
N.Viswanathan de la Universidad de Standford

# Data Generador - transformaciones

## Preprocesado

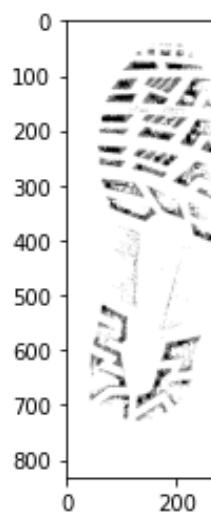


Escala de grises



Threshold binary

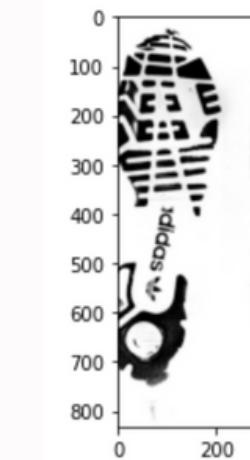
Recortar espacios blancos



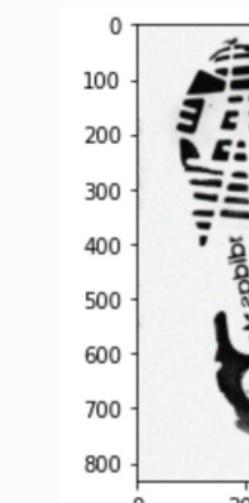
## Aumentación

Previene el sobreentrenamiento añadiendo variaciones en las imágenes de entrenamiento (Transformación aleatoria en cada iteración)

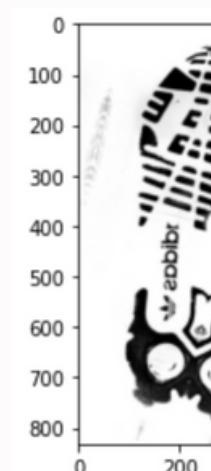
Flip ( voltear )



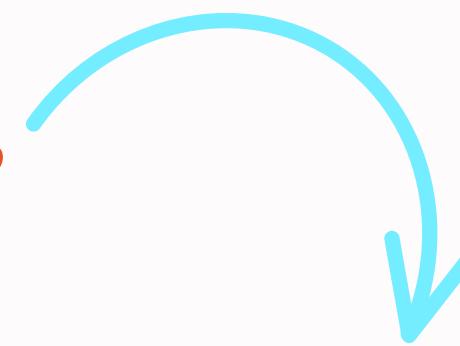
Ruido (distorsión)



Rotate (rotación)



Cuando la IA aprende de manera que solo es aplicable sobre la muestra de *train* y ya no se puede generalizar



# EXPERIMENTOS

Breve análisis de resultados

# Variables

- Eficacia (accuracy)
- Loss (desviación)
- Top 1: porcentaje predicciones correctas
- Top 3: Correcta en las 3 primeras posiciones
- % de similitud: 25%, 50%, 75%
- Matriz de confusión
- Tiempo

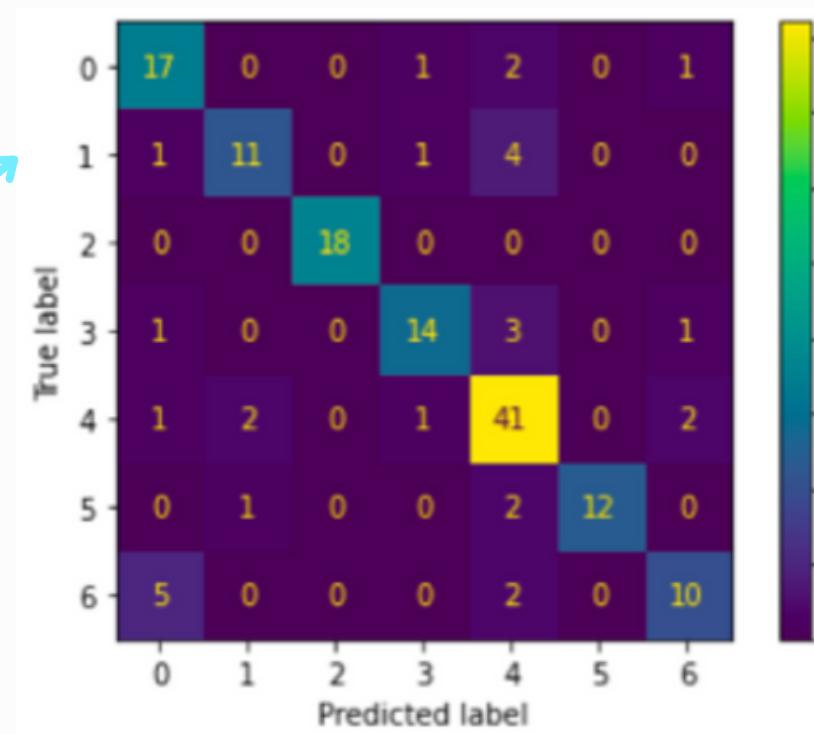


# H1: ¿Qué preprocessado de imágenes funciona mejor?

GBC obtiene mejor accuracy (61% y 79%) en test

Si se utiliza aumentación, el accuracy disminuye, loss aumenta, pero la similitud aumenta.

	epoch:10	test Acc	75% sim
-		0,30	0
A		0,26	0,26
GBC		0,79	0,11
AGBC		0,15	0,15



epoch: 10 y 25

- sin transformación
- A solo aumentación
- G solo escala de grises
- GB blanco y negro
- GBC blanco y negro sin espacios en blanco
- AGBC todas

14 Experimentos en total

## H2: ¿Utilizar más epochs, resulta siempre en mejor resultado?

En general, mayor número de iteraciones, se observa alguna mejora.

Sin aumentación, la eficacia en *train* apenas varia, pero con aumentación aumenta hasta 95%.

20 Experimentos en total

	train Acc	test Acc
-	0,32	0,30
	0,31	0,30
A	0,50	0,04
	0,95	0,15
GBC	0,85	0,68
	0,99	0,81
ABGC	0,30	0,30
	0,47	0,38

epoch = 5  
epoch = 50

epoch: 5, 10, 25 y 50

- - sin transformación
- A solo aumentación
- GB blanco y negro
- GBC blanco y negro sin espacios en blanco
- AGBC todas

# H3: ¿Un mayor tamaño de las imágenes utilizadas, resulta en un mejor resultado?

El tiempo necesario para entrenar correlacionado con el tamaño.

En general, el *accuracy* en test aumenta con el tamaño, pero sobretodo al aplicar ampliación.



	time	train Acc	test Acc	
411x1035	34'	0,78	0,79	GBC
	22'	0,38	0,64	AGBC
208x832	29'	0,97	0,77	
	21'	0,46	0,54	
104x228	13'	0,85	0,67	
	13'	0,37	0,27	

epoch: 10

- GBC blanco y negro sin espacios en blanco
- AGBC todas

Tamaño de imagen:

- 411x1035
- 280x832
- 205x526
- 104x228

# H4: ¿Cuál es el mínimo de muestras por marca aceptable para el modelo?

Todas las muestras -> peor eficacia y similitud

Menos marcas clasificar, tenemos más muestras por marca y mejor la eficacia.

20 Experimentos en total

	marcas	test Acc	75% sim
1	59	0,53 0,15	0,15 0
5	7	0,79 0,15	0 0,15
10	3	0,76 0,29	0,57 0,29

epoch: 10

- - sin transformación
- A solo aumentación
- GBC blanco y negro sin espacios en blanco
- AGBC todas

Mínimo de muestras: 1, 2, 3, 5 y 10

# OTROS EXPERIMENTOS



Autokeras propone un modelo que se sobrentrena.



Utilizar el modelo preentrenado ImageNet (VGG16) con aumentación y transferencia obtiene una eficacia de entrenamiento del 99% pero del 15% en la evaluación.

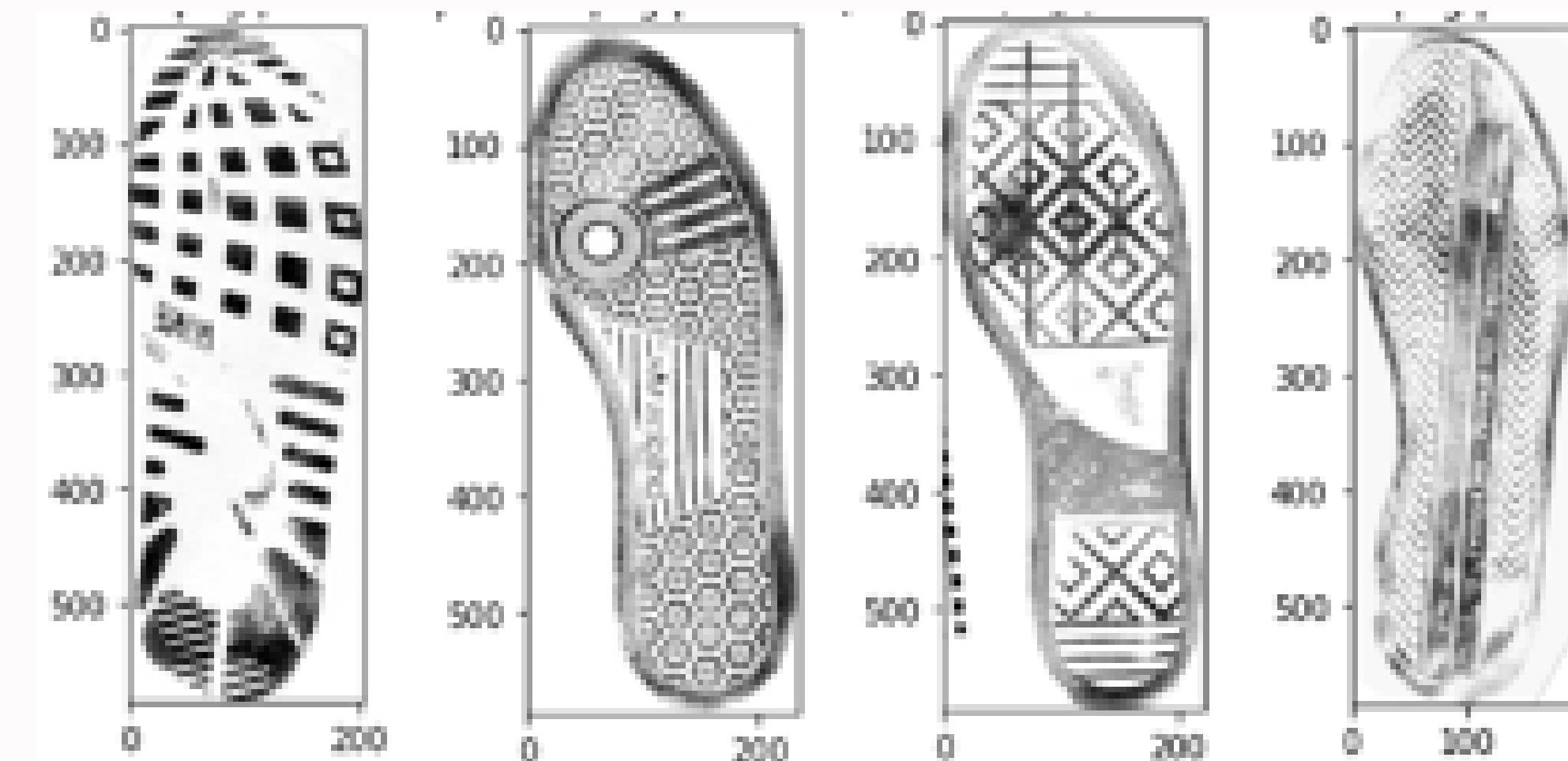


**Variación en las capas del modelo:**

- Sin la capa maxPooling los parámetros aumentan x15 y se triplica el tiempo de ejecución.
- Las capas ocultas colaboran en disminuir el tiempo de ejecución y número de parámetros.
- Dropout ayuda en la eficacia.

# VALIDACIÓN CRUZADA

TRAIN: 100% 2d Footwear  
TEST: FD-300



No satisfactoria, todas las muestras clasificadas como "Nike".  
Red "sobreespecializada": solo funciona en el escenario ideal de  
evaluar imágenes extraídas y procesadas de la misma manera

# OBJETIVO 2

Red neuronal siamesa para  
la búsqueda de imágenes  
similares

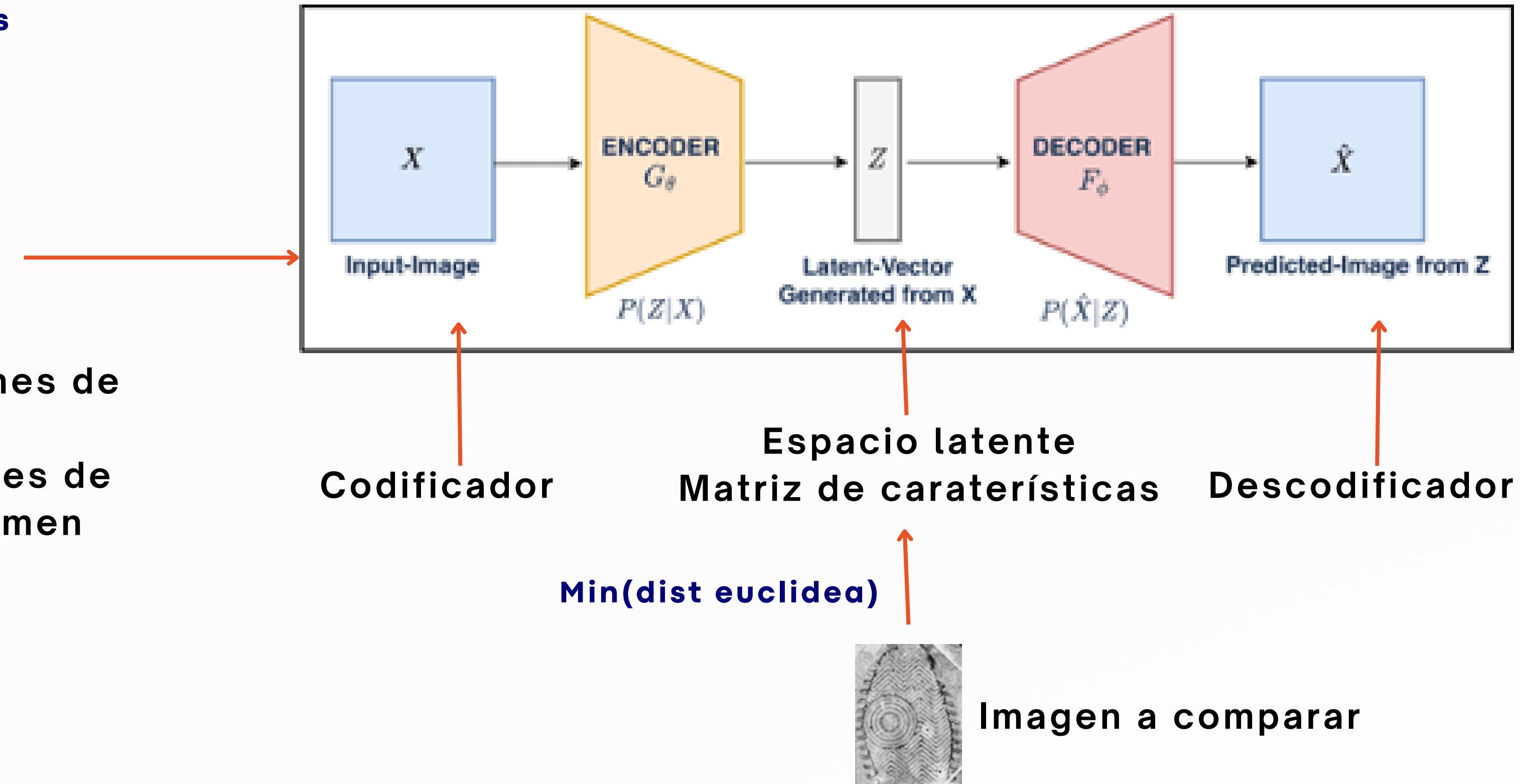
# Red neuronal siamesa

Base de datos

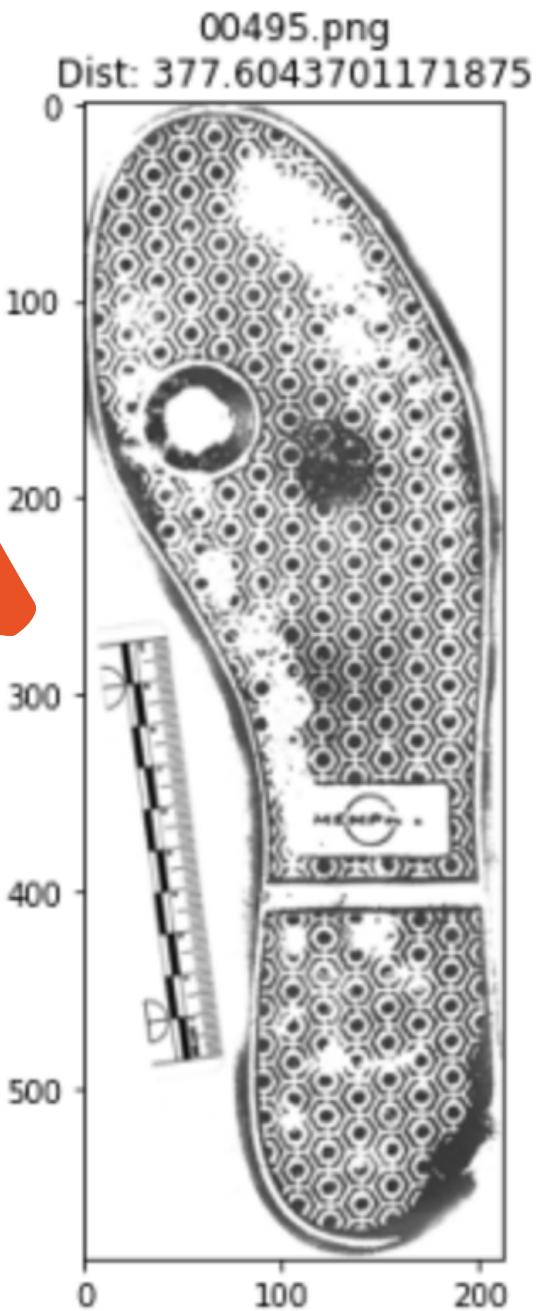
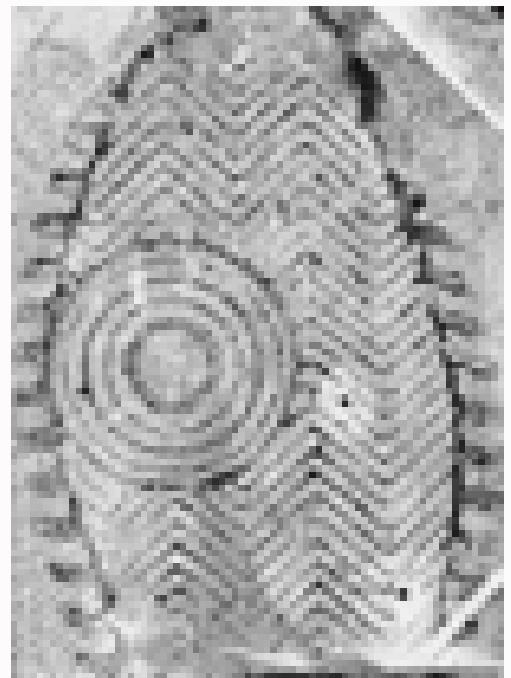
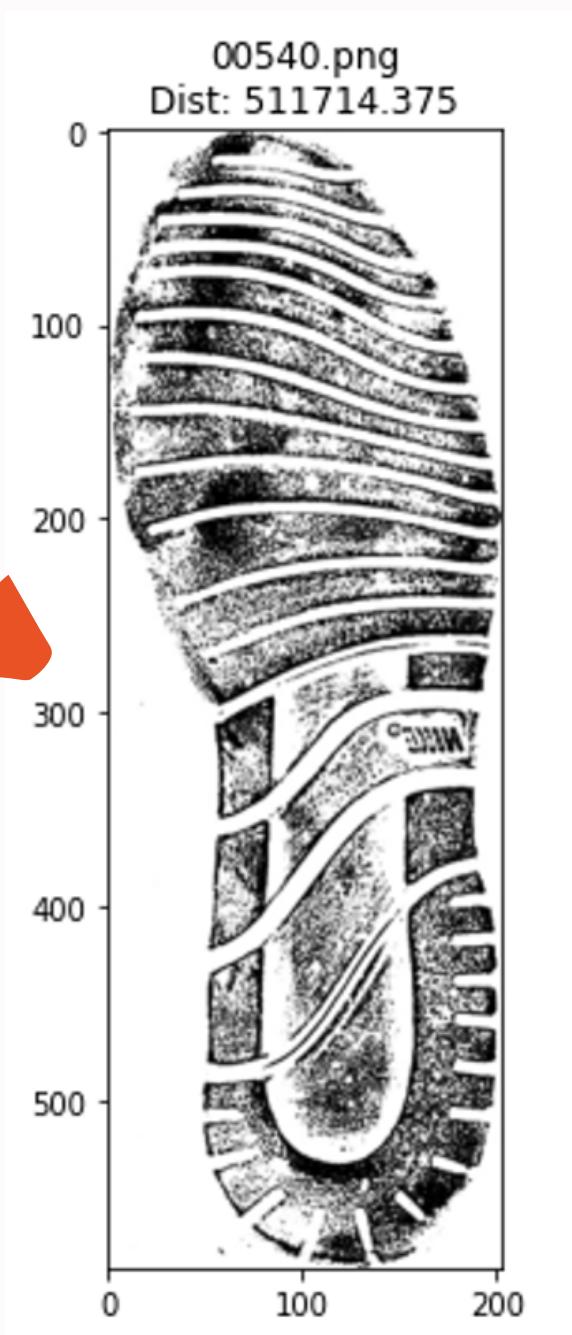
**FD300**



- 1175 imágenes de referencia
- 300 imágenes de escenas crimen



# Resultados red siamesa

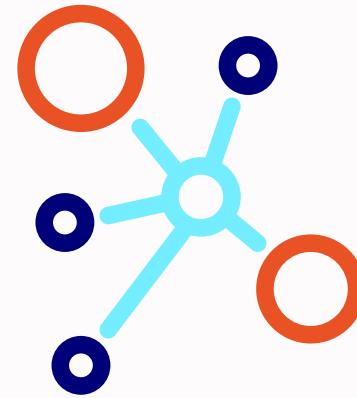


Distancia euclídea

Cierta similitud, sobretodo en **patrones**

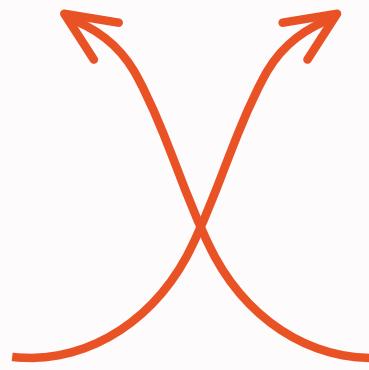
• • • • •  
• • • • •  
• • • • •  
• • • • •

# Conclusión global



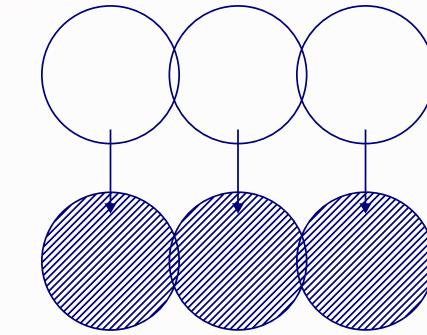
## Red Neuronal Convolucional

- La selección de capas es importante para un buen funcionamiento.
- Transformaciones y aumentación necesarias.
- Vigilar con el tamaño de las imágenes.
- Limitar la clases.



## Validación cruzada

- Modelo con aparente buena eficacia.
- Al aplicar-la se pone a prueba para ver si puede generalizar.
- Resultado insatisfactorio.



## Red neuronal siamesa

- Enfoque académico para consolidar conocimientos.
- Resultado satisfactorio.

# MUCHAS GRACIAS

*"Go in a suit if you commit a crime –  
there are no prints on dress shoes"*  
Srihari (Wired)

**[github.com/laurivsan/TFM\\_AI\\_2023](https://github.com/laurivsan/TFM_AI_2023)**

**[laurivsan@uoc.edu](mailto:laurivsan@uoc.edu)**

