

INSTITUTO NEBRIJA

Reconocimiento de imágenes mediante aprendizaje profundo



Autor: Martín Vilasánchez Freijomil

Tutor: Rubén González Martín

Índice

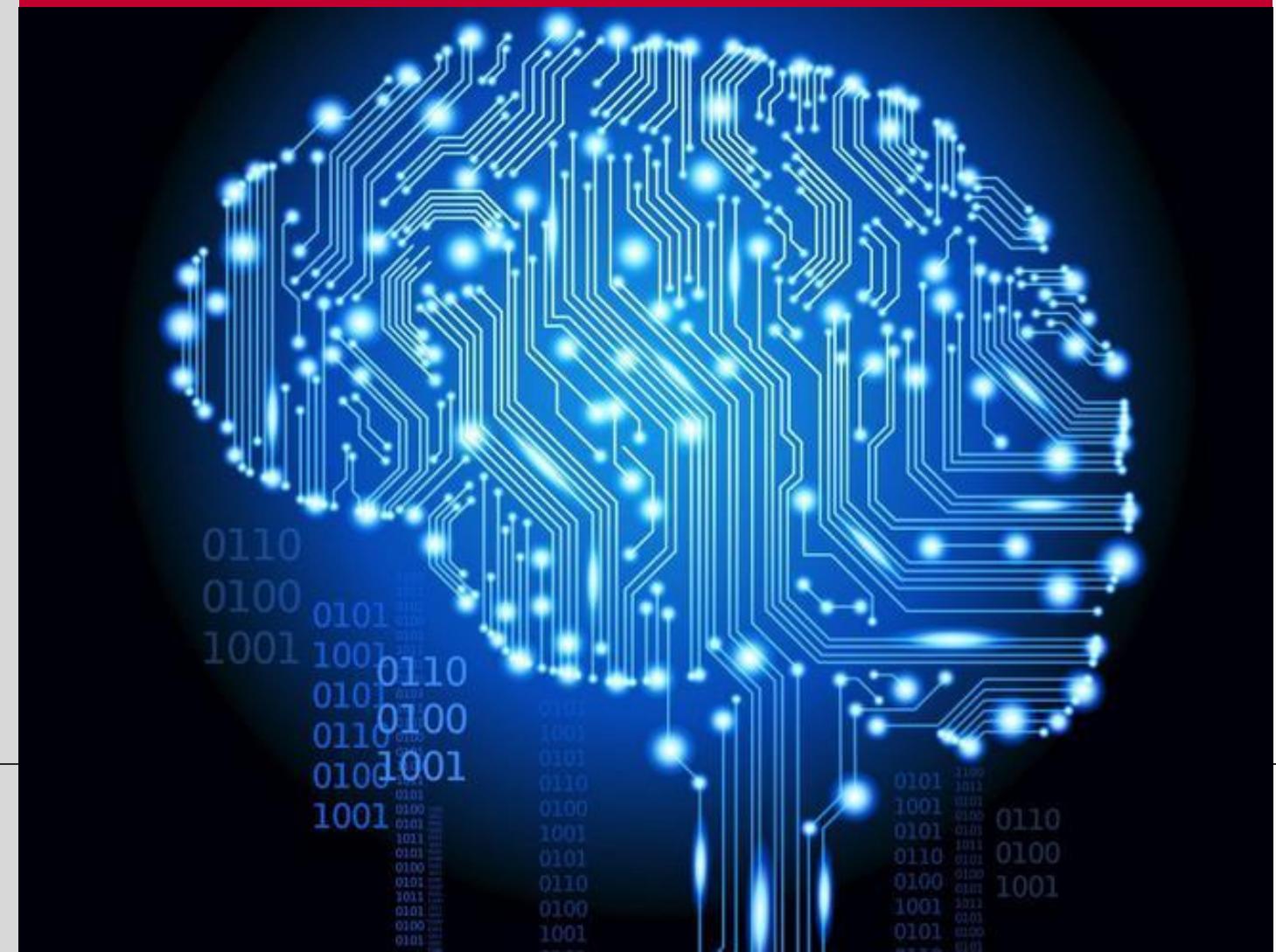
-
- 1. Introducción
 - 2. Objetivos
 - 3. Marco teórico
 - 4. Enfoque y método seguido
 - 5. Resultados
 - 6. Conclusiones
-



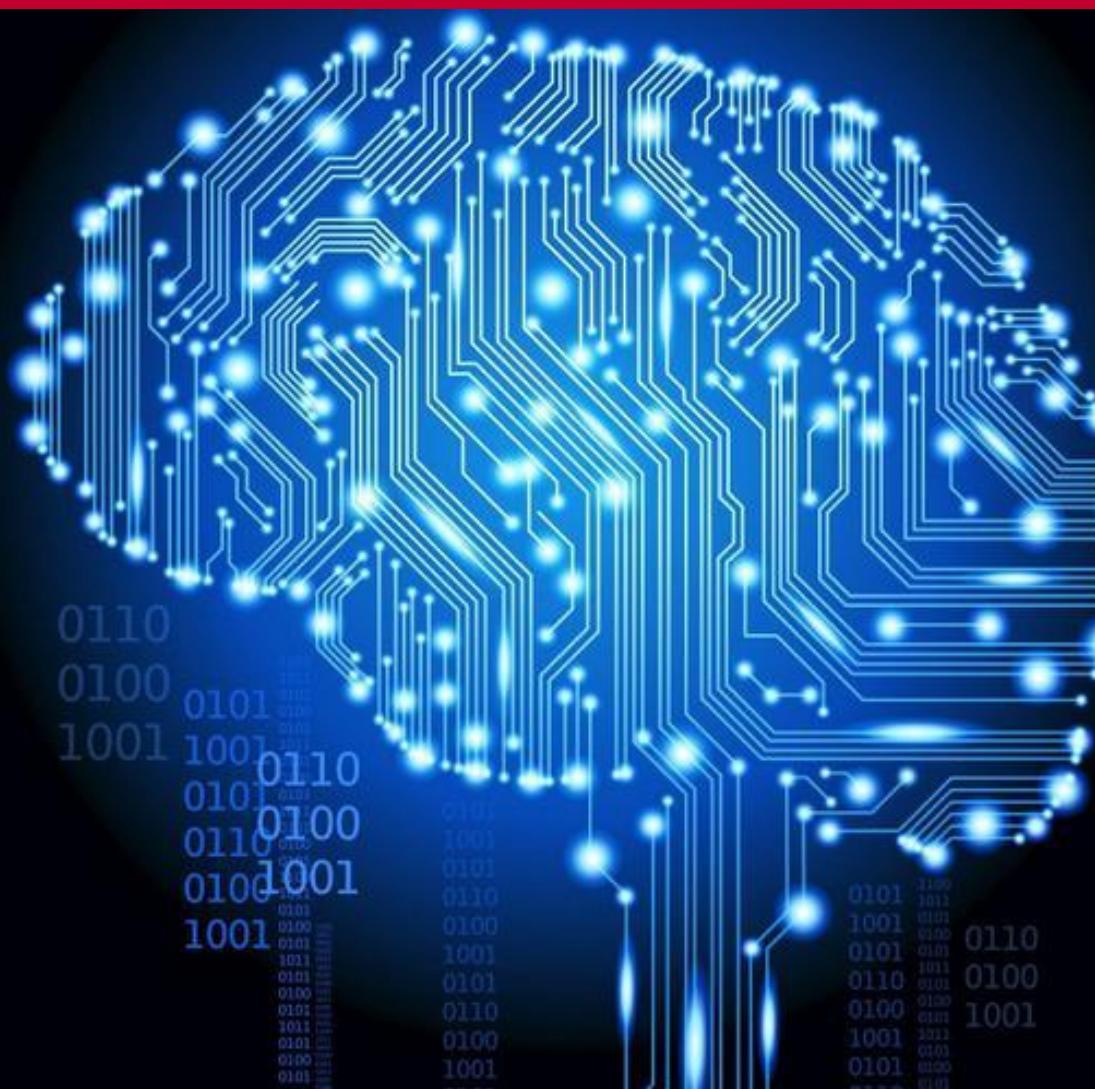
Exploraremos el reconocimiento automático de imágenes, cuya finalidad es construir un clasificador que sea capaz de distinguir entre diez mil clases de seres vivos.

- Haremos uso de redes neuronales convolucionales (CNN) para la construcción de nuestro modelo.
- El conjunto de datos que utilizaremos para entrenar y validar nuestro modelo ha sido descargado de “**iNat Challenge 2021 - FGVC8**”.

Introducción



Objetivos



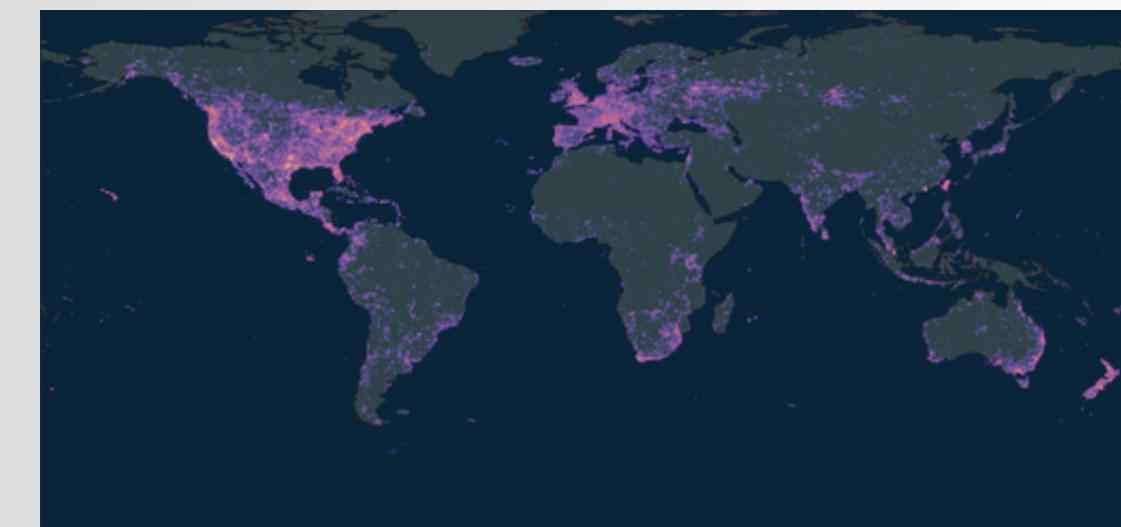
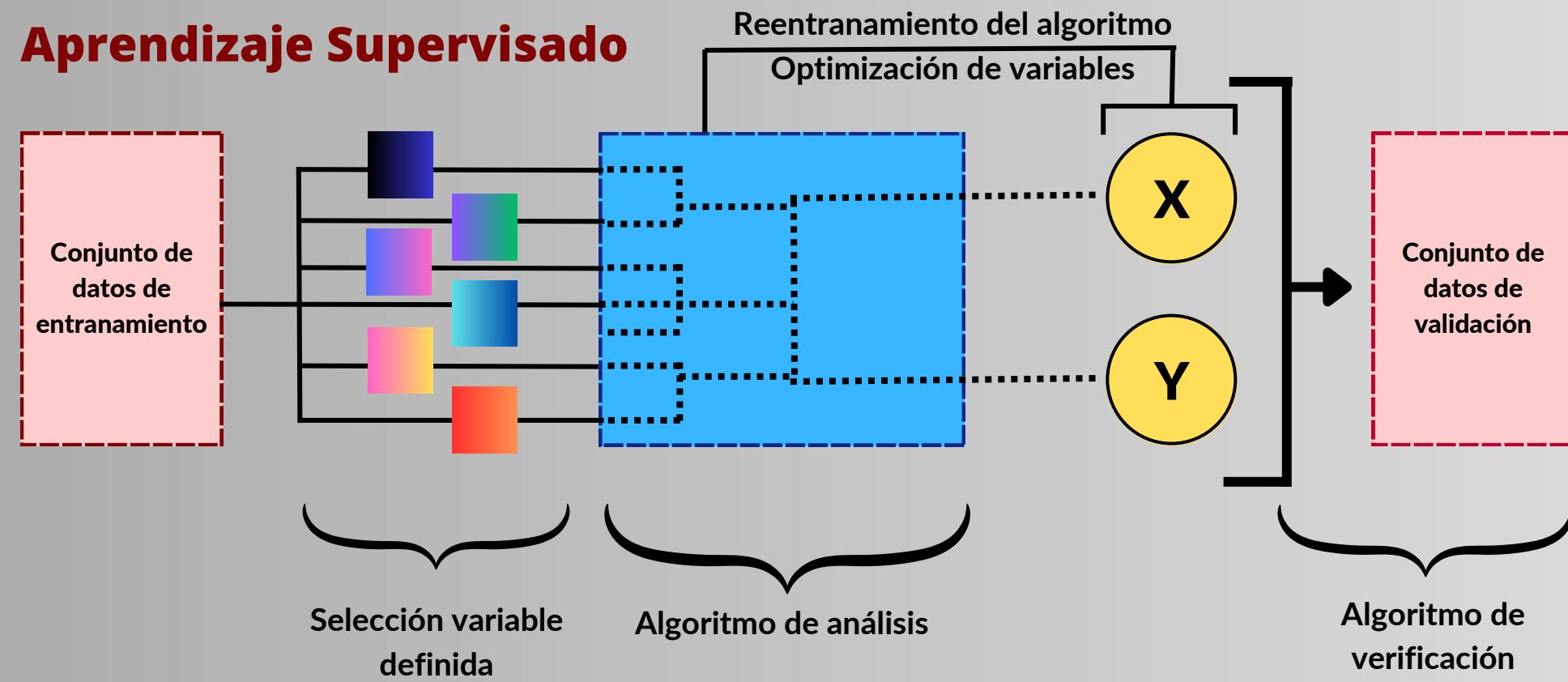
Implementar un clasificador de imágenes, capaz de reconocer y clasificar imágenes de seres vivos: animales, hongos y plantas.

- 1** Desarrollar un clasificador basado en redes neuronales convolucionales.
Comprender el funcionamiento de diferentes modelos de redes neuronales convolucionales y entrenarlos.
- 2** Establecer criterios de selección del modelo.
Analizar los resultados de los entrenamientos de los diferentes modelos.
- 3** Utilizar frameworks y librerías relacionadas con el aprendizaje profundo.
Familiarizarse con el manejo de las herramientas con mayor implantación dentro del aprendizaje profundo.
- 4** Hacer una prueba de concepto.
Desarrollar una aplicación web

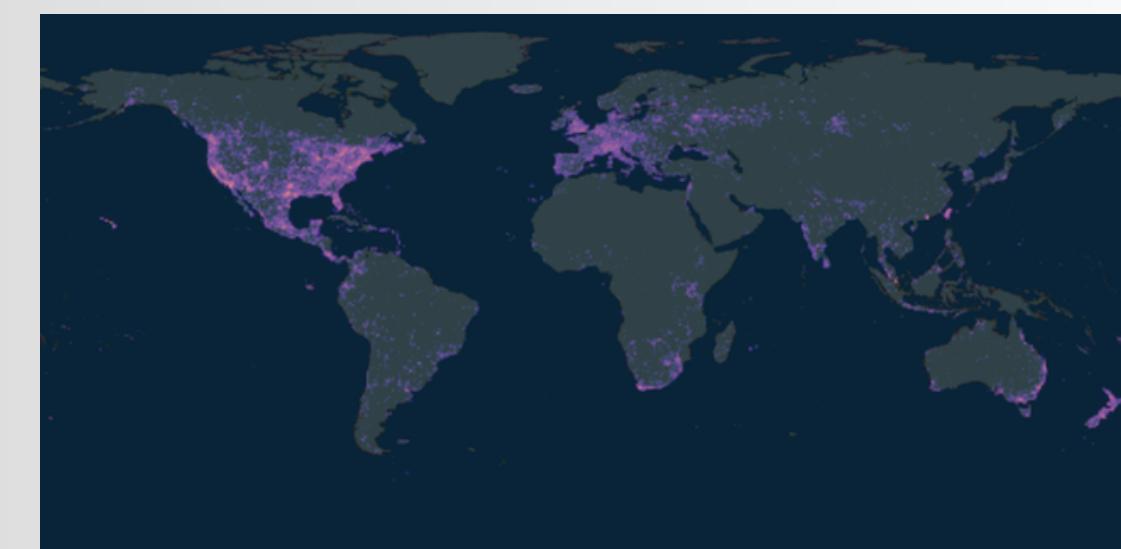
Marco teórico

Arthur Samuel (**1959**, IBM): acuñó el término machine learning o aprendizaje automático como el campo de estudio que confiere a los ordenadores la capacidad de aprender sin ser programados explícitamente.

Aprendizaje Supervisado

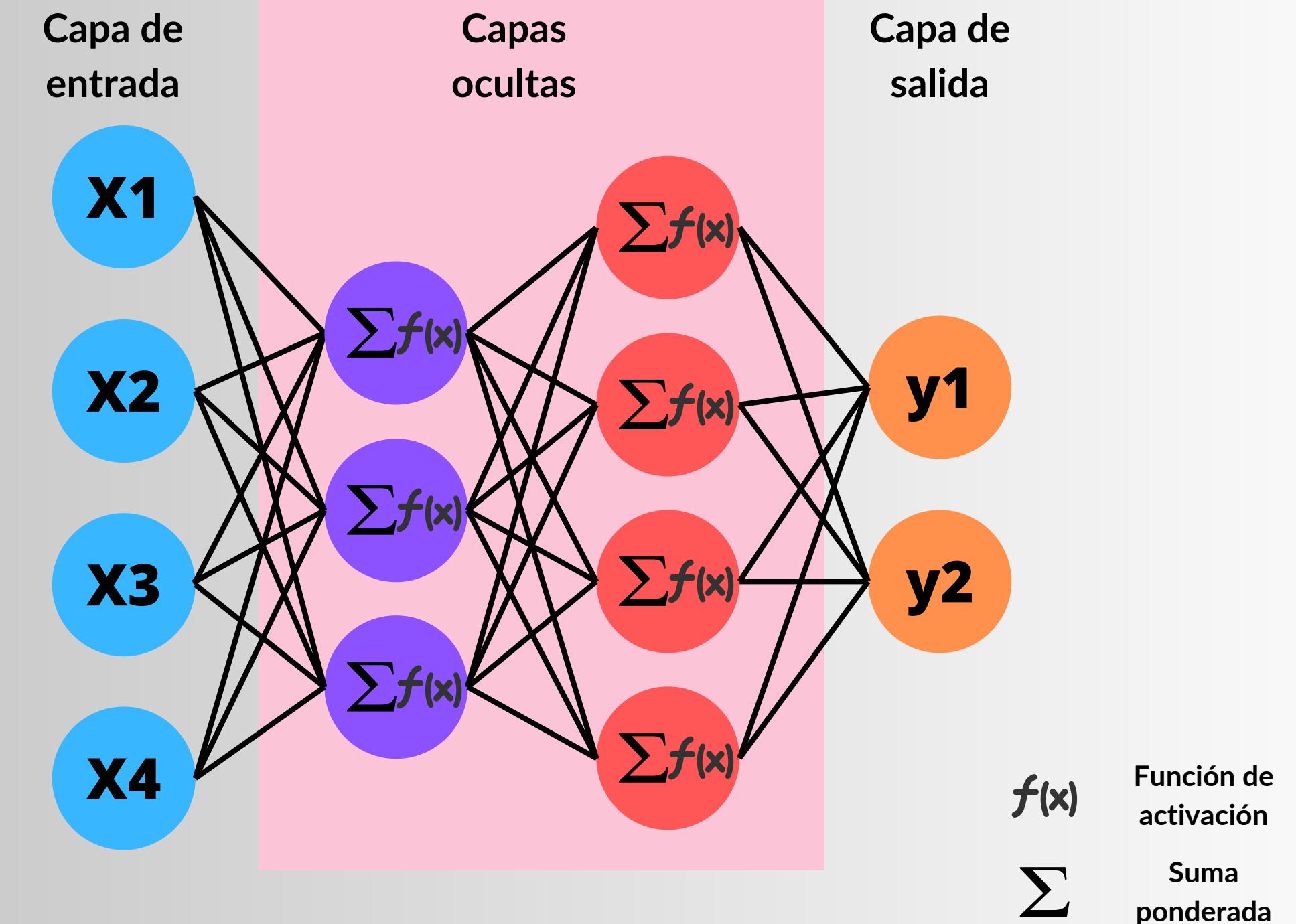
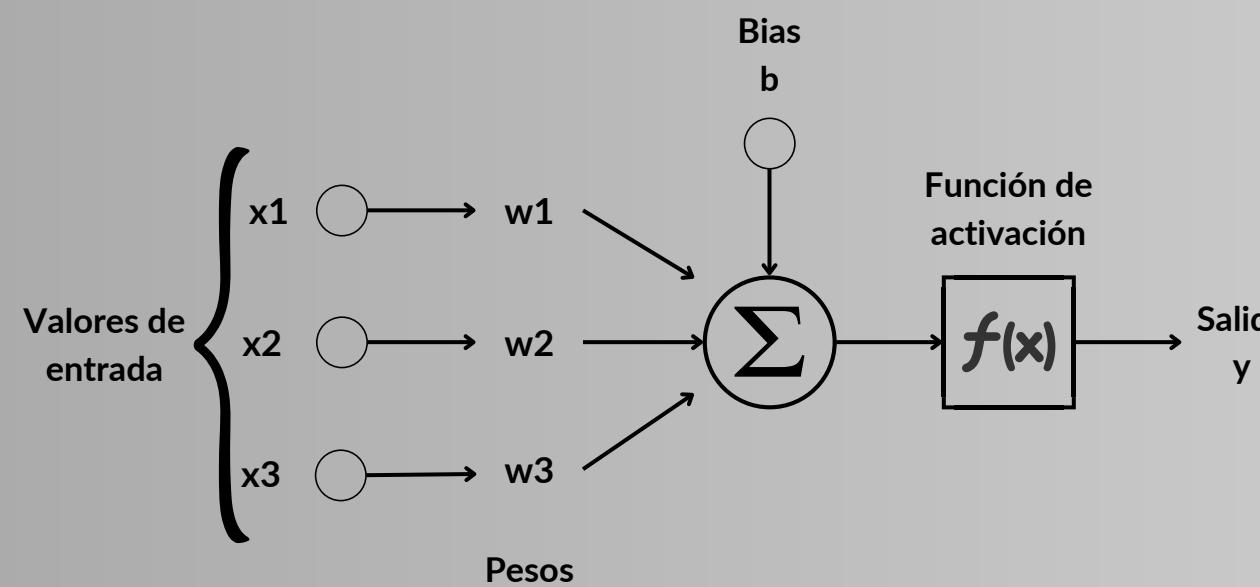


Conjunto de entrenamiento



Conjunto de validación

Redes Neuronales



Redes Neuronales Convolucionales

Mapa de segmentación
de entrada

50	50	50	50	50	50	50
50	230	230	230	230	230	50
50	230	50	50	50	50	50
50	230	230	230	230	50	50
50	230	50	50	50	50	50
50	230	230	230	230	230	50
50	50	50	50	50	50	50

X

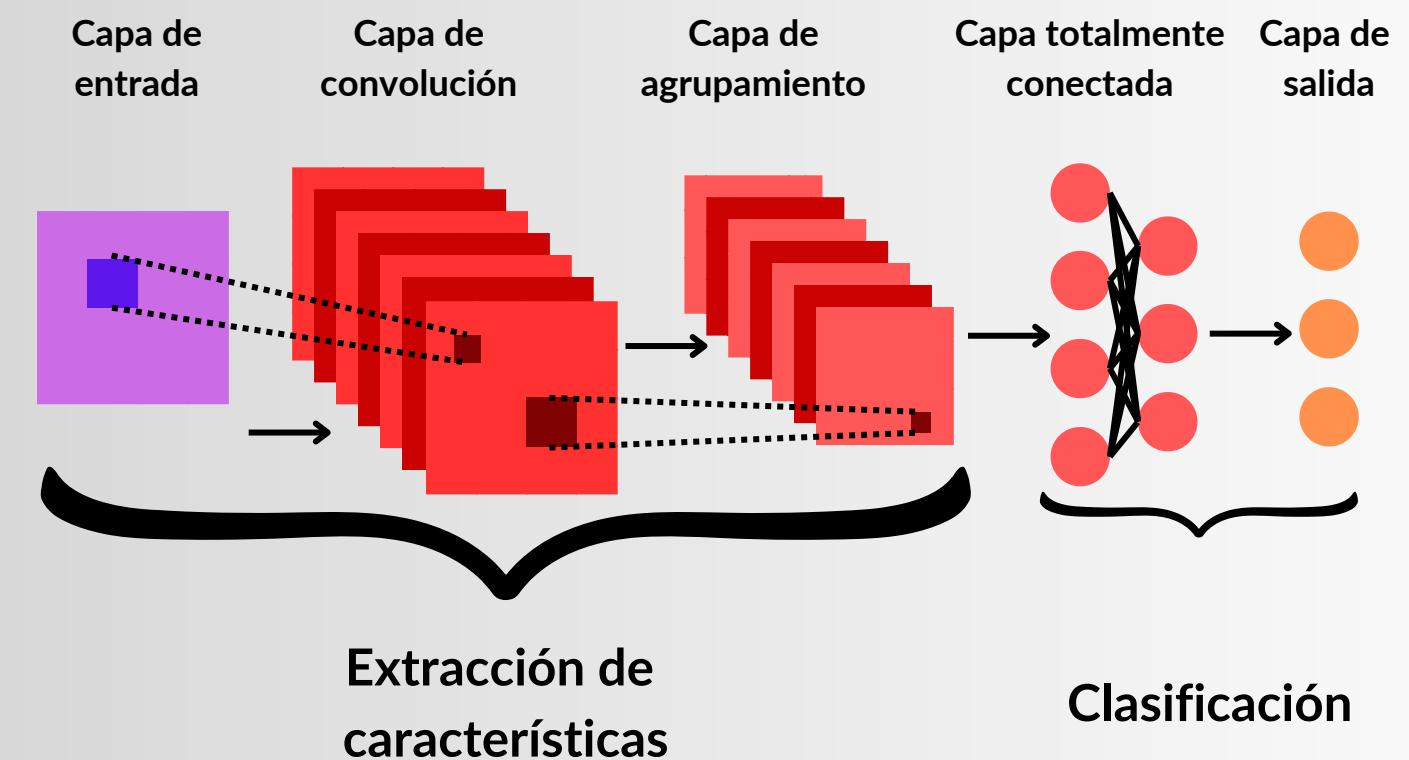
$$\begin{matrix} 0 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & 0 \end{matrix} \quad = \quad \begin{matrix} 0 & 50 & 0 \\ 50 & -460 & 230 \\ 0 & 230 & 0 \end{matrix}$$

KERNEL

Mapa de segmentación
de salida

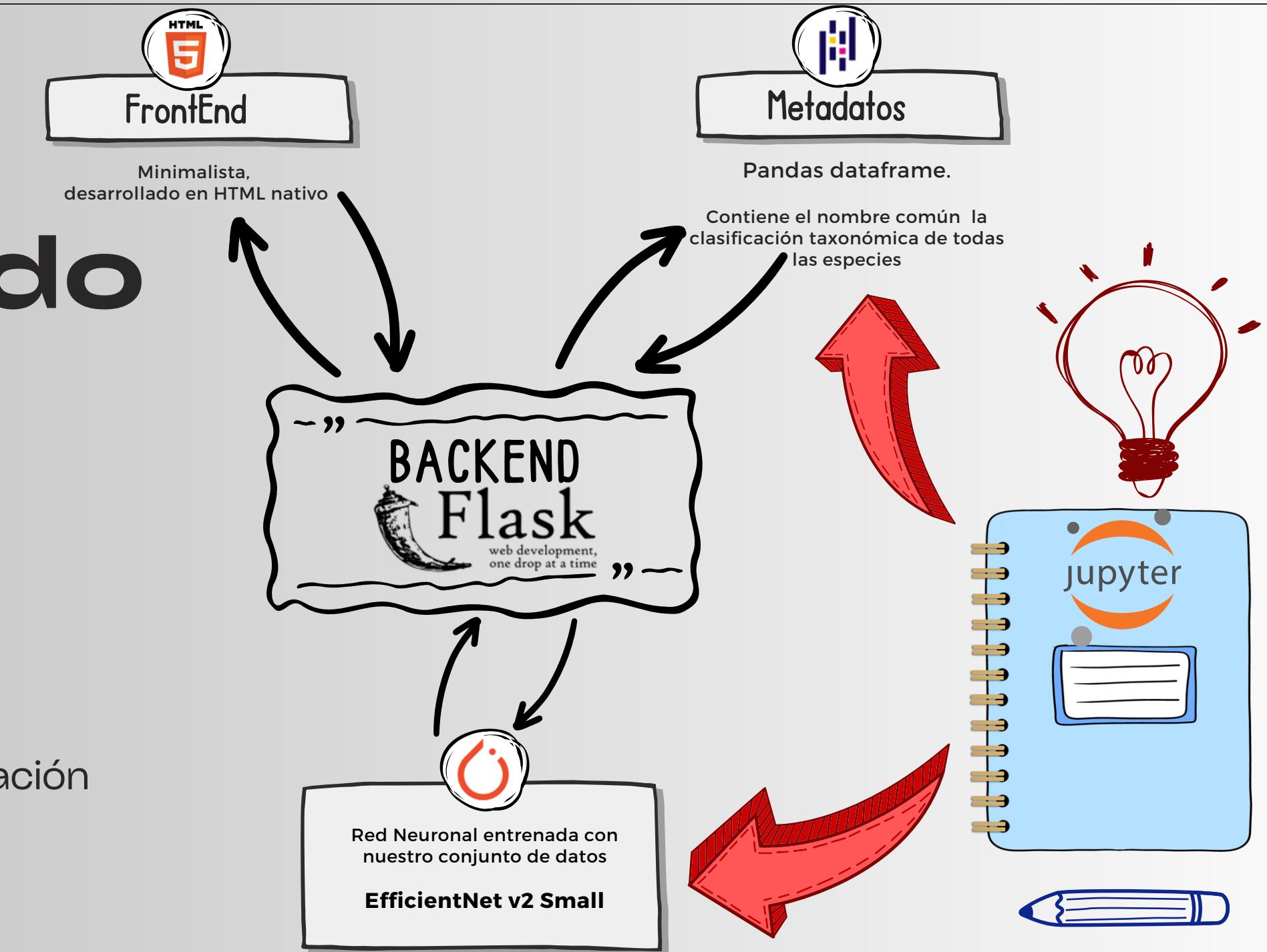
100			

Σ



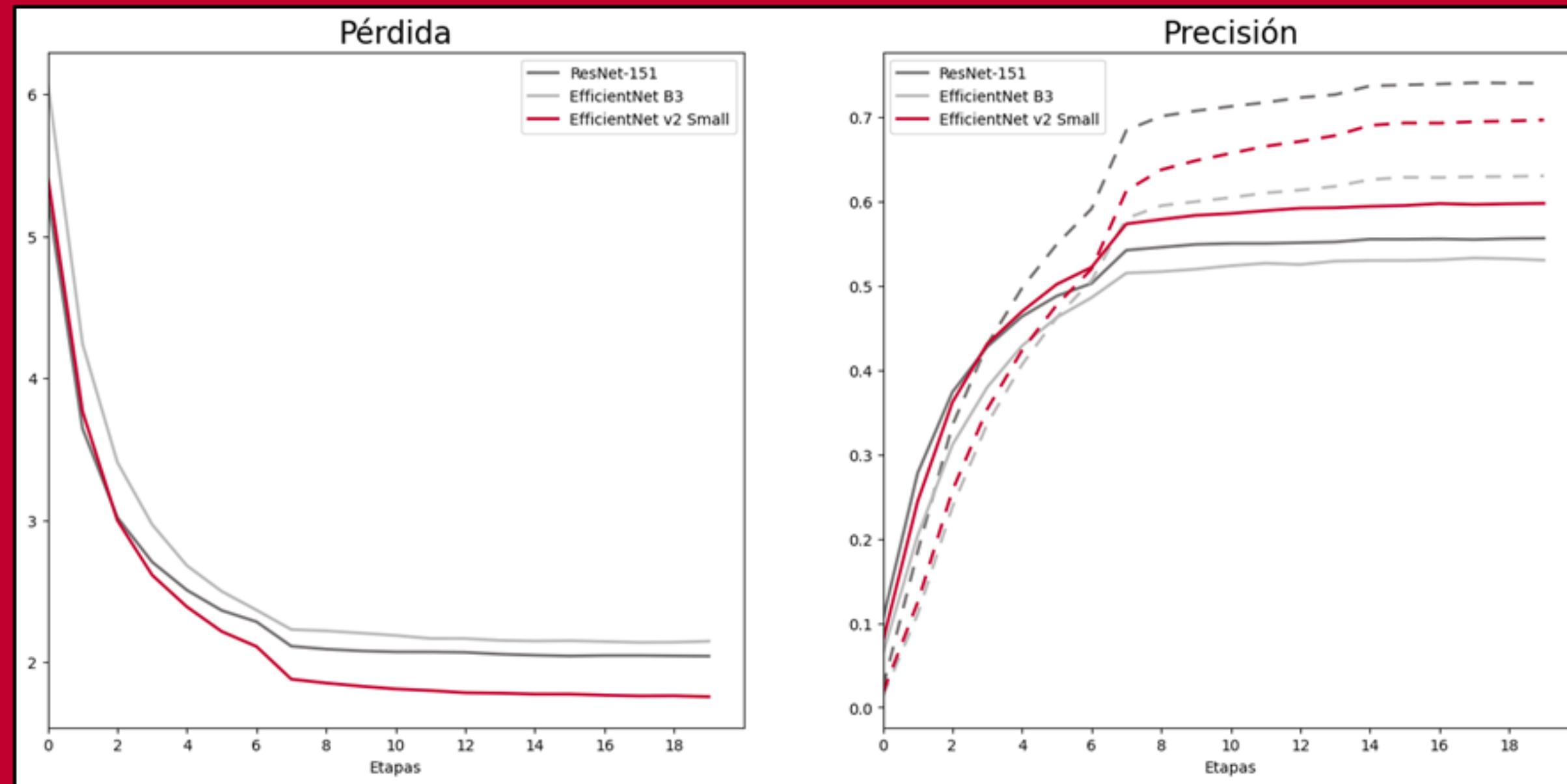
Enfoque y método seguido

1. Estudio de las redes CNN y sus diferentes arquitecturas.
2. Análisis del conjunto de datos.
3. Pre-procesado del conjunto de datos e implementación de los modelos.
4. Entrenamiento y selección del modelo.
5. Prueba de concepto: desarrollo de una aplicación web.



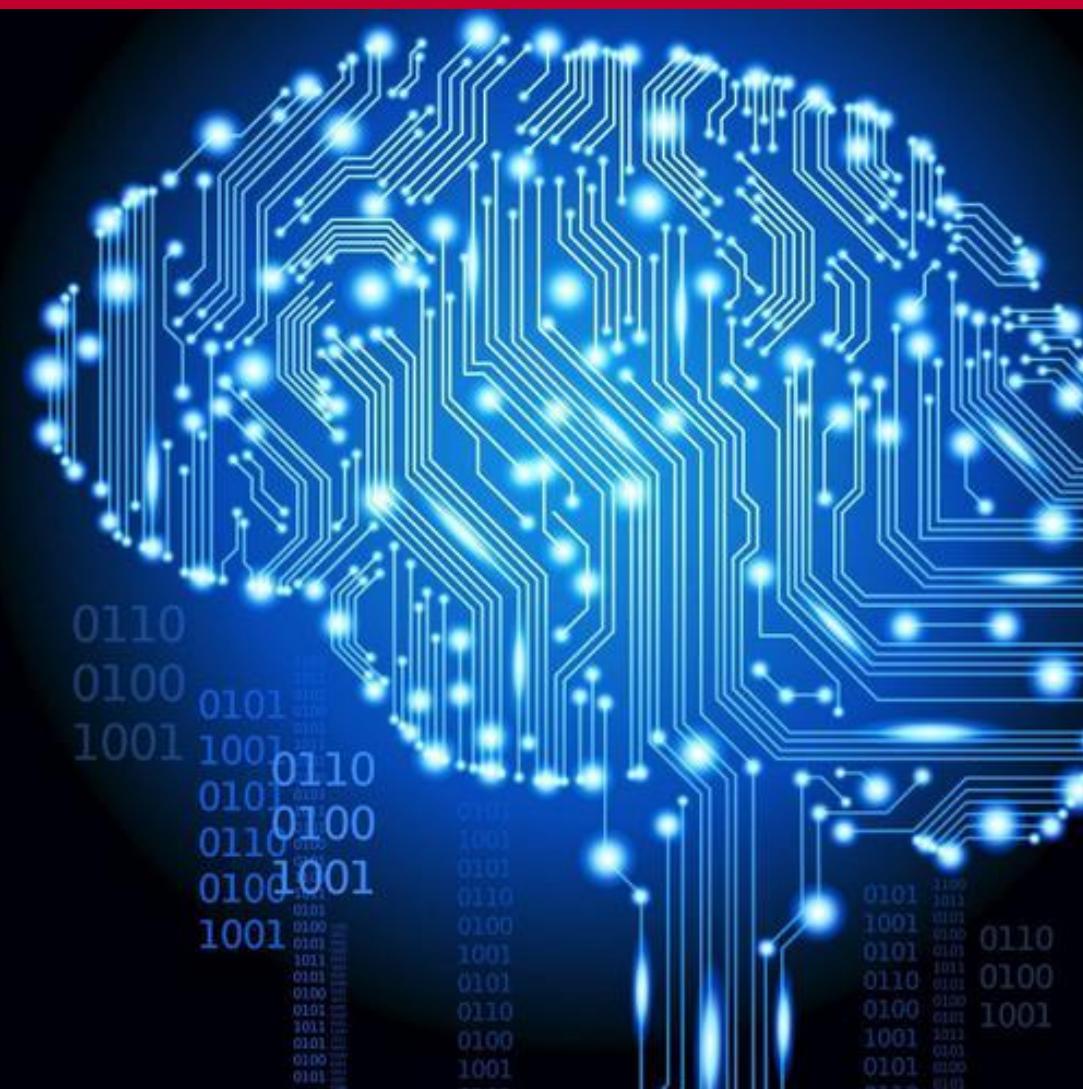
Resultados

Modelo	Valor de pérdida	porcentaje de precisión
ResNet-50	2,0433	55,66%
EfficientNet B3	2,1402	53,31%
EfficientNet v2 Small	1,7577	59,82%



Fuente:
Elaboración propia

Conclusiones



1

Elección del entorno de trabajo

El entrenamiento de la mayoría de los modelos se nos ha ido por encima de las 24 horas. Realizando 25 etapas de entrenamiento. Deberíamos de emplear un equipo con mas memoria RAM y dos GPUs.

2

No ha funcionado el aprendizaje por transferencia.

En el entrenamiento de los diferentes modelos, cuando hemos aplicado el aprendizaje por transferencia ha dado peores resultados.

Conclusiones

¿porqué no ha funcionado el aprendizaje por transferencia en nuestro caso?

- ImageNet está formado por 1.000 categorías diferentes, frente a las 10.000 categorías de nuestro modelo.
- La similitud visual entre categorías diferentes



*Pearl-bordered
Fritillary*



*Pacific
Fritillary*

Reconocimiento de imágenes de flora y fauna mediante CNN

Reiniciar herramienta



Seleccione una imagen de su ordenador

Seleccionar archivo Ninguno archivo selec.



Subir imagen

Análisis

Con una precisión del 21.03 '%'

El Nombre común es: Pearl-bordered Fritillary

El reino es: Animalia

El phylum es: Arthropoda

La clase es: Insecta

El orden es: Lepidoptera

La familia es: Nymphalidae

El género es: Boloria

La especie es: euphrosyne

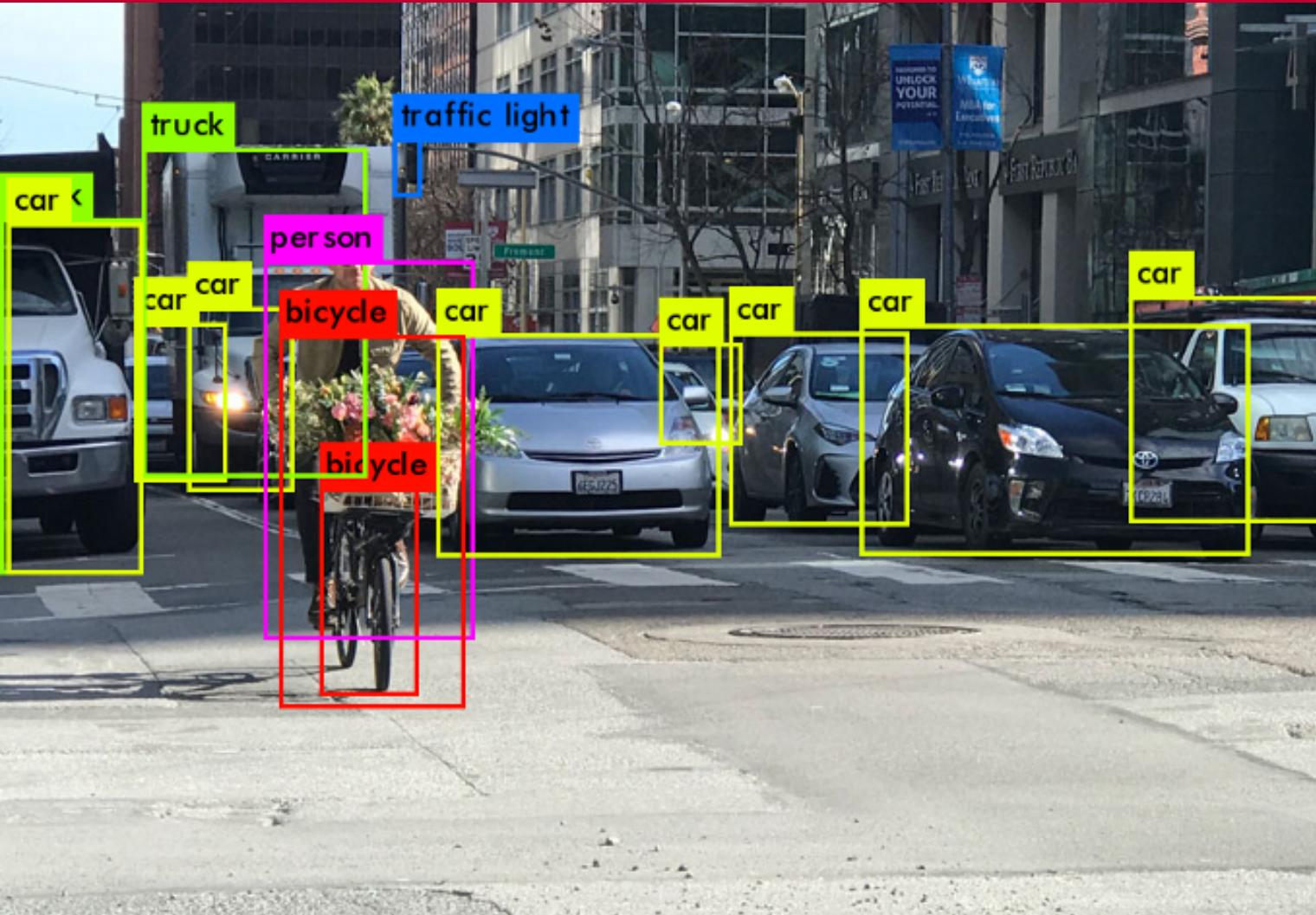


Implementaciones a futuro

1

Entrenar otros modelos.

YOLO (You Only Look Once). Se compone de 137 capas convolucionales preentrenadas y existen herramientas de código abierto que nos permiten adaptar el entrenamiento de las últimas 25 capas





Muchas gracias

JUNIO



2023

