Analizando el Impacto de Modelos Preentrenados con Deep Learning en Tareas de Visión Artificial

Autor: RUBEN GONZALEZ VELASCO
Tutor(a): EMILIO SERRANO FERNANDEZ

Inputs

Classes

cat, dog, bird

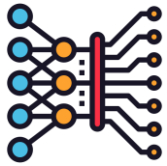
Zero-Shot Image Classification Model

Output

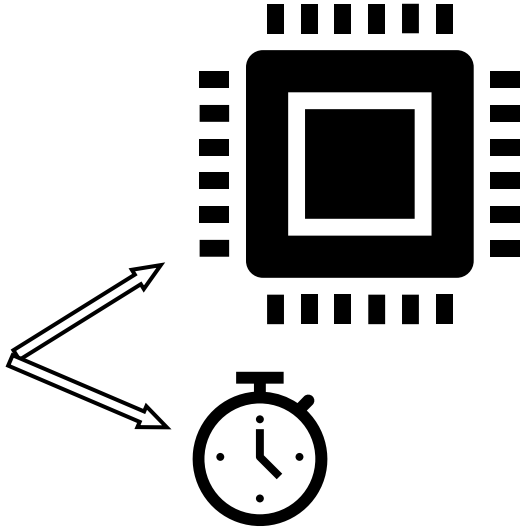
Cat	0.664
Dog	0.329
Bird	0.008

Hugging Face

DEEP LEARNING



1. Recopilación y preproceso de datos
2. Selección del modelo
3. Construcción y Entrenamiento del modelo
4. Evaluación del modelo

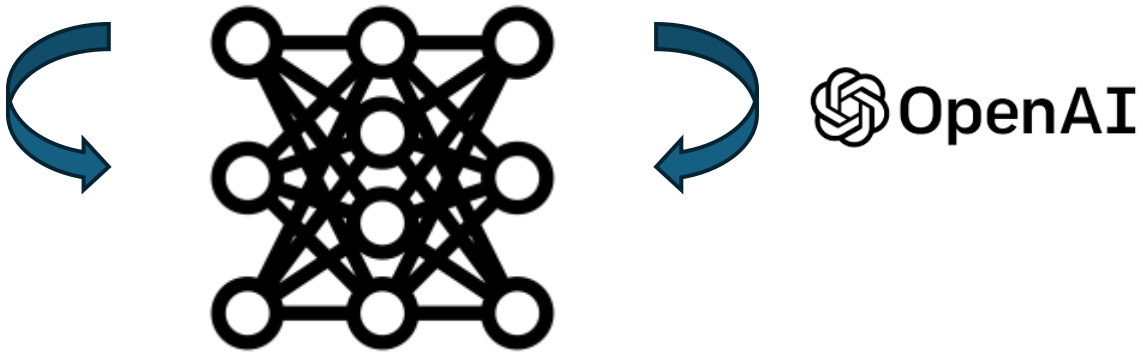


DEEP LEARNING

1. Recopilación y preproceso de datos
2. Selección del modelo
3. Construcción y Entrenamiento del modelo
3. Evaluación del modelo

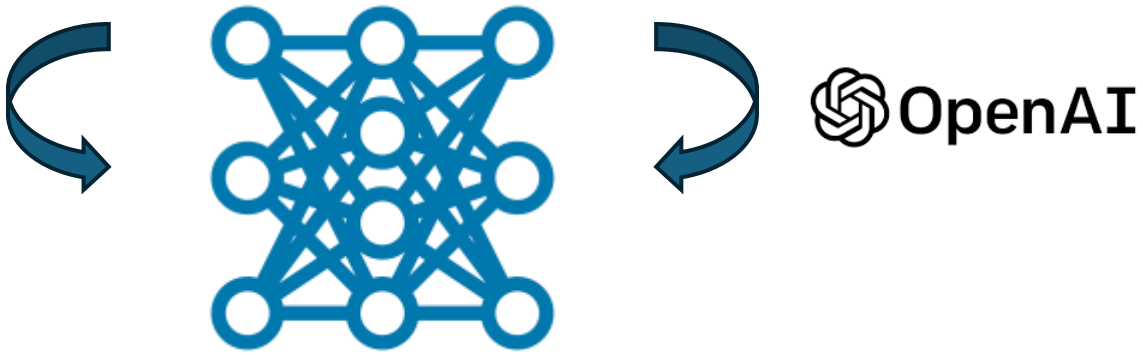
FEW SHOT LEARNING + MODELOS PRE-ENTRENADOS

MODELOS PREENTRENADOS



OpenAI

MODELOS PREENTRENADOS



OpenAI

MODELOS PREENTRENADOS

Imagen nunca vista

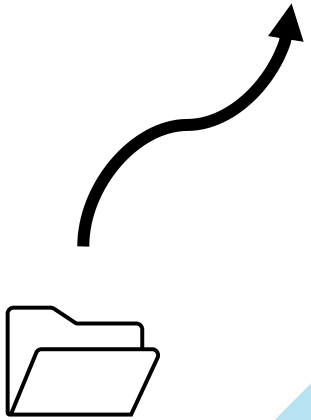
Output

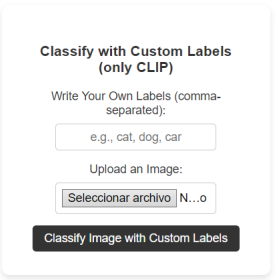
(Imagen)

(Etiquetas)

+

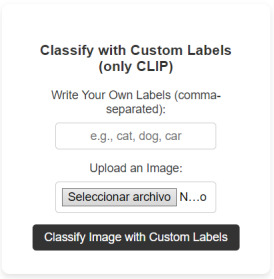
“Shark, Dolphin, Lizard, Giraffe”





Shark

Upload Your Image
Select Label Set:
CIFAR-10
Upload an Image:
Seleccionar archivo
N…o
Classify Image


Classify with Custom Labels (only CLIP)
Write Your Own Labels (comma-separated):
e.g., cat, dog, car
Upload an Image:
Seleccionar archivo
N…o
Classify Image with Custom Labels

Upload Your Image

Select Label Set:

CIFAR-10

Upload an Image:

Seleccionar archivo N…o

Classify Image

Classify with Custom Labels (only CLIP)

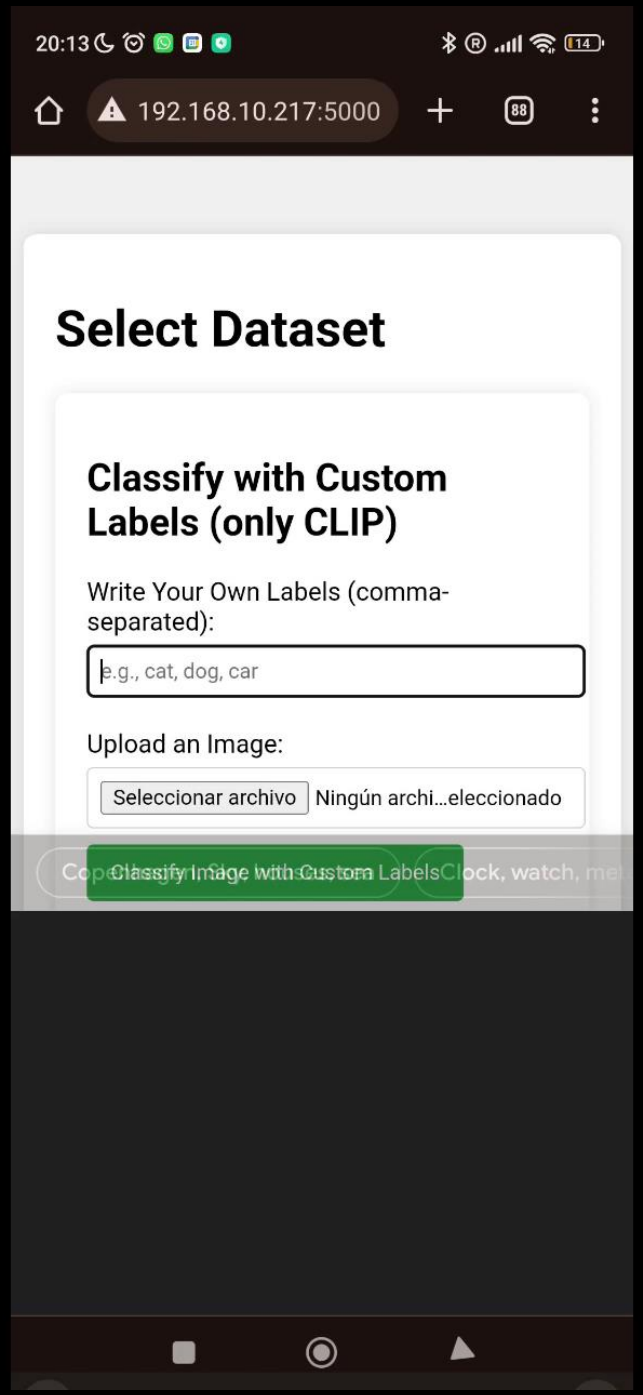
Write Your Own Labels (comma-separated):

e.g., cat, dog, car

Upload an Image:

Seleccionar archivo N…o

Classify Image with Custom Labels


20:13
192.168.10.217:5000
Select Dataset
Classify with Custom Labels (only CLIP)
Write Your Own Labels (comma-separated):
e.g., cat, dog, car
Upload an Image:
Seleccionar archivo
Ningún archi...eleccionado
Classify Image with Custom Labels

IMAGENET
14,197,122 images, 21841 synsets indexed
Home Download Challenges About
Not logged in. Login | Signup
ImageNet Large Scale Visual Recognition Challenge (ILSVRC)
2009
2012
2017
2021
LLMs?
Deep Neural Networks
CNN – Alex Net
Transformers
Attention is All You Need
OpenAI
CLIP
Contrastive
Language-Image
Pre-training
ResNet
VGG
(Visual Geometry Group)

Actual: **truck**

ResNet Predictions:

- dog - 62.21%
- deer - 11.75%
- frog - 9.89%

Classification Time: 1.05 seconds

VGG Predictions:

- deer - 99.85%
- truck - 0.09%
- cat - 0.06%

Classification Time: 0.18 seconds

Clip Predictions:

- truck - 86.43%
- ship - 6.7%
- automobile - 4.23%

Classification Time: 0.18 seconds

Actual: **truck**

ResNet Predictions:

- truck - 85.08%
- horse - 7.78%
- frog - 3.95%

Classification Time: 0.92 seconds

VGG Predictions:

- cat - 99.98%
- horse - 0.02%
- dog - 0.0%

Classification Time: 0.17 seconds

Clip Predictions:

- truck - 98.33%
- automobile - 1.14%
- ship - 0.25%

Classification Time: 0.16 seconds

Actual: **bird**

ResNet Predictions:

- deer - 52.94%
- dog - 37.04%
- truck - 4.75%

Classification Time: 0.9 seconds

VGG Predictions:

- deer - 99.91%
- horse - 0.08%
- airplane - 0.01%

Classification Time: 0.17 seconds

Clip Predictions:

- bird - 87.8%
- frog - 3.83%
- dog - 2.54%

Classification Time: 0.16 seconds

Actual: **horse**

ResNet Predictions:

- horse - 99.7%
- ship - 0.2%
- airplane - 0.07%

Classification Time: 0.9 seconds

VGG Predictions:

- automobile - 99.99%
- bird - 0.01%
- deer - 0.0%

Classification Time: 0.17 seconds

Clip Predictions:

- horse - 47.57%
- bird - 16.12%
- airplane - 12.05%

Classification Time: 0.16 seconds

Actual: **truck**

ResNet Predictions:

- ship - 35.13%
- horse - 31.73%
- cat - 24.35%

Classification Time: 0.94 seconds

VGG Predictions:

- truck - 100.0%
- deer - 0.0%
- frog - 0.0%

Classification Time: 0.18 seconds

Clip Predictions:

- truck - 96.84%
- automobile - 2.62%
- ship - 0.26%

Classification Time: 0.19 seconds

Actual: **dog**

ResNet Predictions:

- airplane - 98.07%
- truck - 1.87%
- dog - 0.02%

Classification Time: 0.91 seconds

VGG Predictions:

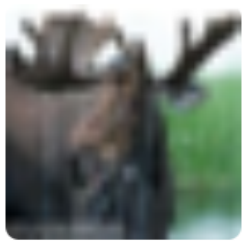
- dog - 100.0%
- automobile - 0.0%
- cat - 0.0%

Classification Time: 0.16 seconds

Clip Predictions:

- dog - 90.29%
- frog - 1.91%
- airplane - 1.67%

Classification Time: 0.16 seconds



Actual: **deer**

ResNet Predictions:

- bird - 88.43%
- frog - 3.82%
- dog - 3.57%

Classification Time: 0.9 seconds

VGG Predictions:

- truck - 100.0%
- automobile - 0.0%
- ship - 0.0%

Classification Time: 0.16 seconds

Clip Predictions:

- deer - 94.29%
- horse - 4.62%
- bird - 0.39%

Classification Time: 0.16 seconds

Actual: **automobile**

ResNet Predictions:

- truck - 82.73%
- bird - 13.68%
- cat - 3.47%

Classification Time: 0.89 seconds

VGG Predictions:

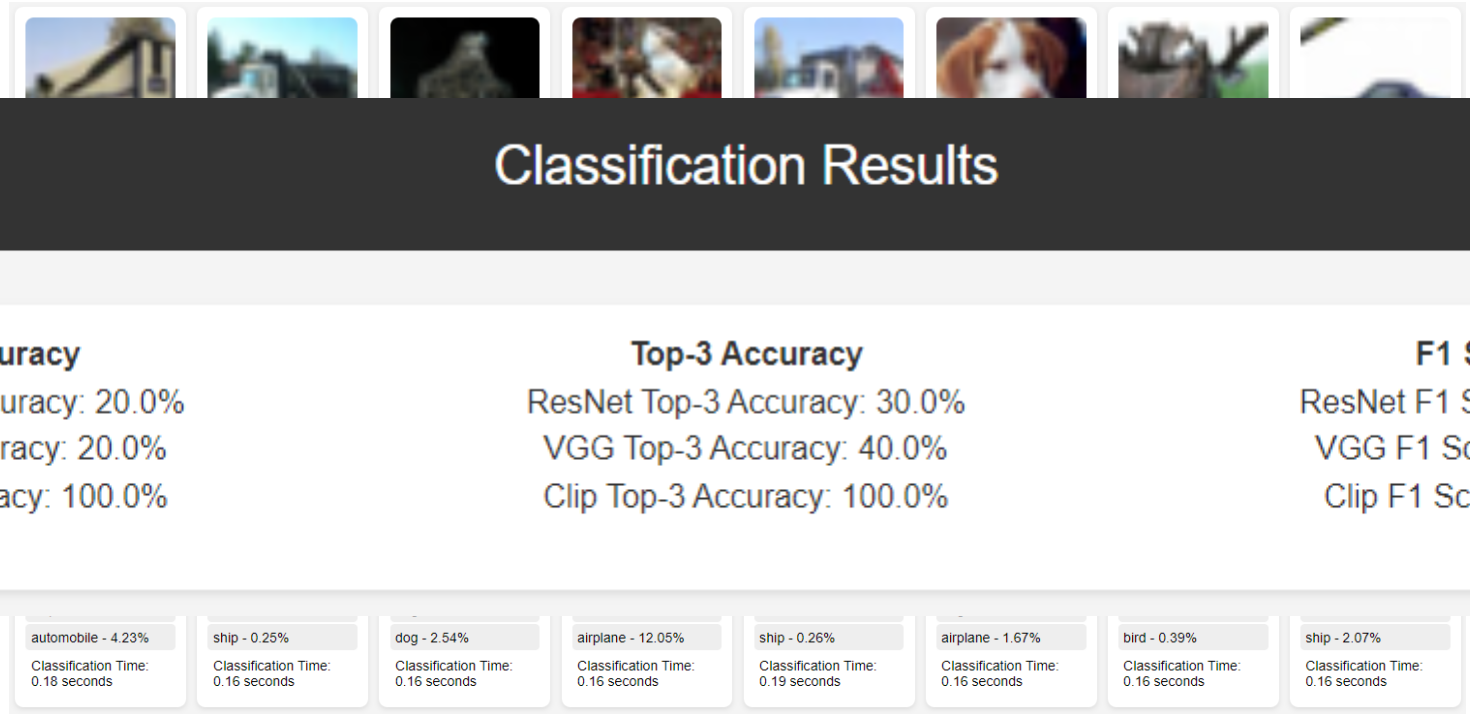
- bird - 100.0%
- horse - 0.0%
- automobile - 0.0%

Classification Time: 0.16 seconds

Clip Predictions:

- automobile - 86.17%
- truck - 8.32%
- ship - 2.07%

Classification Time: 0.16 seconds



Classification Results

Accuracy

ResNet Accuracy: 20.0%

VGG Accuracy: 20.0%

Clip Accuracy: 100.0%

Top-3 Accuracy

ResNet Top-3 Accuracy: 30.0%

VGG Top-3 Accuracy: 40.0%

Clip Top-3 Accuracy: 100.0%

F1 Score

ResNet F1 Score: 11.85%

VGG F1 Score: 12.86%

Clip F1 Score: 100.0%

Select Dataset for Classification

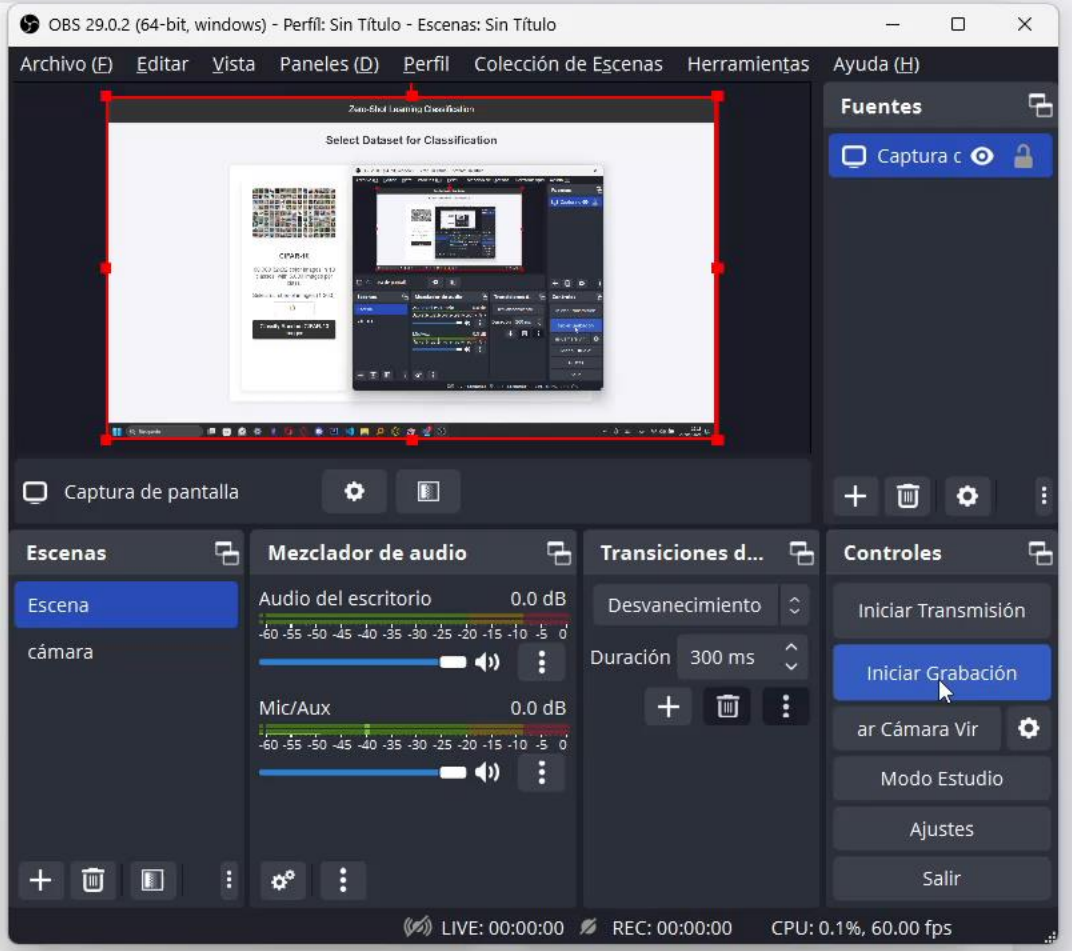
CIFAR-10

60,000 32x32 color images in 10 classes, with 6,000 images per class.

Select number of images (1-200):

10

Classify Random CIFAR-10 Images



Back

Classification Results

Accuracy
ResNet Accuracy: 10.0%
VGG Accuracy: 8.5%
Clip Accuracy: 86.5%

Top-3 Accuracy
ResNet Top-3 Accuracy: 30.5%
VGG Top-3 Accuracy: 28.0%
Clip Top-3 Accuracy: 98.0%

F1 Score
ResNet F1 Score: 9.88%
VGG F1 Score: 8.31%
Clip F1 Score: 86.37%

Actual: **cat** | Actual: **ship** | Actual: **bird** | Actual: **airplane** | Actual: **dog** | Actual: **deer** | Actual: **bird** | Actual: **frog** | Actual: **ship**

200 imágenes de CIFAR 10

Accuracy
ResNet Accuracy: 10.0%
VGG Accuracy: 8.5%
Clip Accuracy: 86.5%

Top-3 Accuracy
ResNet Top-3 Accuracy: 30.5%
VGG Top-3 Accuracy: 28.0%
Clip Top-3 Accuracy: 98.0%

F1 Score
ResNet F1 Score: 9.88%
VGG F1 Score: 8.31%
Clip F1 Score: 86.37%

Back

Classification Results

Accuracy
ResNet Accuracy: 10.0%
VGG Accuracy: 8.0%
Clip Accuracy: 6.0%

Top-3 Accuracy
ResNet Top-3 Accuracy: 29.5%
VGG Top-3 Accuracy: 29.5%
Clip Top-3 Accuracy: 29.5%

F1 Score
ResNet F1 Score: 9.73%
VGG F1 Score: 7.84%
Clip F1 Score: 4.36%

Actual: **Dress** | Actual: **Trouser** | Actual: **Pullover** | Actual: **T-shirt/top** | Actual: **Pullover** | Actual: **Sandal** | Actual: **Dress** | Actual: **Ankle boot** | Actual: **Bag**

200 imágenes de FASHION MNIST

Accuracy
ResNet Accuracy: 10.0%
VGG Accuracy: 8.0%
Clip Accuracy: 6.0%

Top-3 Accuracy
ResNet Top-3 Accuracy: 29.5%
VGG Top-3 Accuracy: 29.5%
Clip Top-3 Accuracy: 29.5%

F1 Score
ResNet F1 Score: 9.73%
VGG F1 Score: 7.84%
Clip F1 Score: 4.36%