

Clasificación de Imágenes: animals10

Ana Valentina López Chacón

Reconocimiento de Formas y Aprendizaje Automático, 2024 - 2025
MIARFID, UPV

Octubre, 2024

Descripción del problema

Se busca desarrollar un clasificador de imágenes capaz de distinguir entre 10 animales diferentes. Para esto se implementaran técnicas desde Redes Neuronales Sencillas y Búsqueda de Hiperparámetros hasta Redes Convolucionales, Transfer Learning y Fine-tuning.

Motivación

- Establece un balance entre la complejidad suficiente para trabajar con Redes Sencillas y Convolucionales hasta modelos pre-entrenados, permitiendo ver un nivel de evolución entre estas.

Análisis descriptivo de los datos

El dataset animals10 [1] fue creado por Alessio Corrado en 2019 para su tesis. Las imágenes fueron recolectadas de *google images* y han sido validadas manualmente. Además, contiene data de baja calidad para añadir realismo.

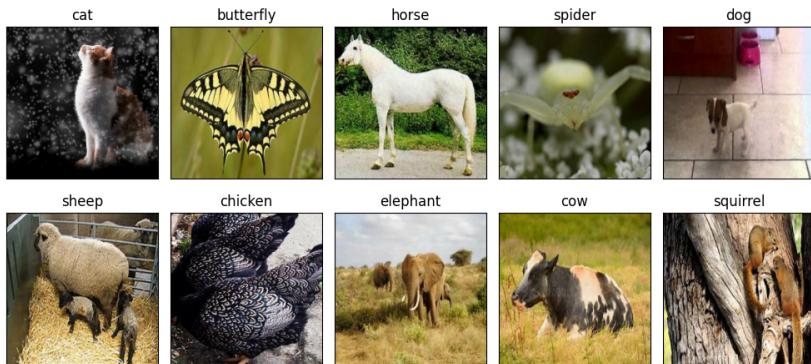


Figura: Gráfica de una Imagen por Clase.

Análisis descriptivo de los datos

- El dataset contiene 26.180 archivos (636 MB) de 10 animales: dog, cat, horse, spyder, butterfly, chicken, sheep, cow, squirrel, elephant.
- La carga del dataset se realizó utilizando el método `image_dataset_from_directory()`.



Resultados Previos

El estado del arte de los proyectos utilizando este dataset está bastante avanzado, solo en Kaggle ha sido utilizado para más de 200 proyectos públicos.

Referencia	Modelo Base	Precisión en test
[2]	VGG16	95 %
[3]	ResNet50	93,56 %
[4]	EfficientNetB7	97,17 %

Resultados con Redes y Ajuste de Hiperparámetros

	Modificación de Red	Hiperparámetros	Precisión en test
MLP básico	4096, ReLU		36,6 %
KerasTuner	4096, ReLU		37,1 %
	256, ReLU		
	64, ReLU		
Ajuste LR y		$lr = 0,001$	37,19 %
BS		$BS = 128$	
ReduceOnPlateau		factor = 0,2585	41,47 %
		Paciencia = 5	

Resultados con CNNs

	Modificación de Red	Hiperparámetros	Precisión en test
Convolucionales	3 Conv 3 MaxPooling	filters = 32	67,47 %
Regularización	kernel_regularizer_l2	Penalidad = 0.02 Dropout = 0.1	58,47 %

Transfer Learning y Fine-tuning

	Red	Hiperparámetros	Modelo Base	Precisión en test
Transfer Learning	512, ReLU Dropout(0.5)	lr = 0.001 BS = 128 factor = 0.2585 patience = 5	EfficientNetB0	97,61 %
Fine-Tuning	Ultimas Capas	lr = 0.00001	ResNet50 EfficientNetB0 ResNet50	96,73 % 97,57 % 96,79 %

Transfer Learning, Fine-tuning y Aumento de Datos

	Red	Hiperparámetros	Modelo Base	Precisión en test
Transfer Learning	Rotation Translation Zoom	$f_aug = 0.04$	EfficientNetB0	97,5 %
Fine-Tuning	Ultimas Capas	$lr = 0.00001$	ResNet50	96,89 %
			EfficientNetB0	97,75 %
			ResNet50	96,89 %

Reporte de Clasificación

	Precision	Recall	F1-Score	Support
dog	0.98	0.98	0.98	947
horse	0.98	0.96	0.97	522
elephant	0.99	0.99	0.99	273
butterfly	0.99	0.97	0.98	429
chicken	0.99	0.99	0.99	593
cat	0.98	0.98	0.98	331
cow	0.92	0.95	0.94	399
sheep	0.95	0.94	0.95	382
spider	0.99	0.99	0.99	979
squirrel	0.99	0.98	0.99	380
accuracy			0.98	5235
macro avg	0.98	0.97	0.98	5235
weighted avg	0.98	0.98	0.98	5235


Conclusiones

- Antes del Transfer Learning, el mejor resultado se da con capas convolucionales sin embargo con mucho overfitting, al agregar la regularización la precisión baja pero con menos sobreentrenamiento.
- El dataset animals10 no es muy robusto y requiere una extracción de características más eficiente con arquitecturas más ligeras, lo cual permite que en general la EfficientNetB0 obtenga mejores resultados que la ResNet50.
- El mejor resultado de precisión en test se obtiene haciendo Fine-tuning con la EfficientNetB0 al hacer aumento de datos con un 97.75 %.
- No siempre el modelo con más parámetros va a otorgar los mejores resultados.


Referencias


 *Animals-10*. (2019, 12 diciembre). Kaggle.

<https://www.kaggle.com/datasets/alessiocrrado99/animals10>

 Marcus-Deans. (2021, 23 marzo). *Animal10-VGG-Classification: CNN Network based on VGG16 for classification of 10 types of animal images*. GitHub.

<https://github.com/marcus-deans/Animal10-VGG-Classification>

 Abdulbasitniazi. (2022, 14 agosto). *ResNet50FromScratch EDA*. Kaggle. <https://www.kaggle.com/code/abdulbasitniazi/resnet50fromscratch-eda#7>

 Vencerlanz. (2023, 8 abril). *Animal Image Classification using EfficientNetB7*. Kaggle. <https://www.kaggle.com/code/vencerlanz09/animal-image-classification-using-efficientnetb7/notebook>