

# CLASIFICADOR DE RAZAS DE PERROS

## VISION TRANSFORMER & TRANSFER LEARNING DE RED NEURONAL CONVOLUCIONAL

Jose Luis Tavera

Universidad de los Andes

Abril 9, 2024

# Tabla de Contenidos

## 1 Problema y Relevancia

- Problema
- Relevancia

## 2 Revisión de Literatura

- Breakthrough Conventional Based Approach for Dog Breed Classification Using CNN with Transfer Learning
- Dog Breed Classification Using CNN
- Exploring Vision Transformers for Fine-grained Classification

## 3 Dataset

## 4 Plan de Trabajo

## 5 Resultados Esperados

# Problema y Relevancia

## Problema

Desde la aparición de las redes neuronales convolucionales y otros métodos de aprendizaje profundo para la detección de imágenes, uno de sus principales objetivos ha sido la clasificación de seres vivos, abarcando desde plantas hasta animales. Esto se debe a los múltiples casos de uso que tienen en la protección y preservación de estas especies.

## Relevancia

Según el informe de la American Humane Association del año 2022, solo en los Estados Unidos hay más de 135 millones de perros y gatos como mascotas. Alarmantemente, más de un tercio de estos animales se pierden durante su vida, y el 80% de ellos nunca logra ser encontrado.

# Breakthrough Conventional Based Approach for Dog Breed Classification Using CNN with Transfer Learning

## Autores

Punyanuch Borwarnginn, Kittikhun Thongkanchorn, Sarattha Kanchanapreechakorn, Worapan Kusakunniran

## Dataset

El conjunto de datos de perros con partes de Columbia, que contiene 8,351 imágenes de perros de 133 razas.

## Método

- 1 Enfoques basados en métodos convencionales: Patrón binario local (LBP), Histograma de gradientes orientados (HOG).
- 2 Enfoque basado en aprendizaje profundo: Redes neuronales convolucionales (CNN) con transfer learning (utilizando la red Inception V3).

# Breakthrough Conventional Based Approach for Dog Breed Classification Using CNN with Transfer Learning

## Resultados

- 1 El modelo CNN reentrenado logró una precisión del 96.75% en la clasificación de razas de perros.
- 2 Los métodos convencionales como HOG lograron una precisión del 79.25%.
- 3 La CNN con transfer learning superó a los métodos convencionales.
- 4 La evaluación se extendió a 133 razas utilizando InceptionV3, MobileNetV2 y NASNet, logrando tasas de precisión de hasta el 91%.

# Dog Breed Classification Using CNN

## Autores

Sandra Varghese y Remya S Departamento de Ingeniería Informática y de Ingeniería, Colegio de Ingeniería Saintgits Kottukulam Hills, Pathamuttom, Kottayam, India

## Dataset

No se especifica un dataset específico, pero se menciona el uso de imágenes de perros para entrenar y probar el modelo de clasificación de razas de perros.

## Método

- 1 Utilización de una Convolutional Neural Network (CNN) para clasificar diferentes razas de perros.
- 2 Empleo de estrategias innovadoras en deep learning, incluyendo redes neuronales convolucionales y transfer learning.

# Dog Breed Classification Using CNN

## Resultados

- 1 El modelo de CNN propuesto logra una precisión del 93.53% y 90.86% en dos datasets diferentes.
- 2 La transferencia de aprendizaje (transfer learning) se utiliza para mejorar la precisión del modelo.
- 3 Se discute que la CNN reentrenada supera significativamente a construir el modelo desde cero.
- 4 Se plantean mejoras potenciales, como el uso de más datos, ajuste del modelo y aumento de datos para mejorar la precisión del modelo de clasificación de razas de perros.

# Exploring Vision Transformers for Fine-grained Classification

## Autores

Marcos V. Conde (Universidad de Valladolid) & Kerem Turgutlu (University of San Francisco)

## Dataset

Utilizan varios datasets de fine grained models entre esos el Stanford Dogs con 120 clases, 12,000 imágenes en entrenamiento y 8,580 imágenes de prueba

## Método

- 1 Se utilizó un enfoque basado en Vision Transformers (ViT) para la clasificación de imágenes detalladas.
- 2 Se propuso un marco multi-escenario multi-escala que utiliza ViT para localizar y reconocer regiones informativas en las imágenes.
- 3 Se hizo data augmentation para mejorar las capacidades del modelo.



# Exploring Vision Transformers for Fine-grained Classification

## Resultados

Método	Backbone	Precisión (%)
MaxEnt [5]	DenseNet-161	83.6
FDL [REF]	DenseNet-161	84.9
RA-CNN [6]	VGG-19	87.3
Cross-X [15]	ResNet-50	88.9
API-Net [26]	ResNet-101	90.3
ViT [4]	ViT-B/16	91.7
Ours	ViT-B/16	93.2

# Stanford Dogs Dataset

## Descripción

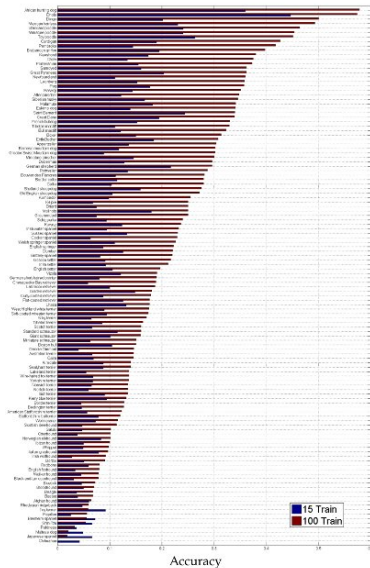
El conjunto de datos Stanford Dogs contiene imágenes de 120 razas de perros de todo el mundo. Este conjunto de datos se ha construido utilizando imágenes y anotaciones de ImageNet para la tarea de categorización de imágenes detalladas. Contenidos de este conjunto de datos:

- **Número de categorías:**120
- **Número de imágenes:** 20,580
- **Anotaciones:** Etiquetas de clase, Cuadros delimitadores

# Stanford Dogs Dataset



## Stanford Dogs Dataset



# Plan de Trabajo

- ➊ **Preparación del conjunto de datos:** Descarga y organiza las imágenes de las 120 razas de perros del conjunto de datos Stanford Dogs, junto con las anotaciones necesarias.
- ➋ **Cargar un modelo pre-entrenado:** Selecciona una arquitectura pre-entrenada (Inception V3) y cargarla utilizando TensorFlow.
- ➌ **Modificar la última capa:** Reemplaza la última capa con una nueva capa de salida para las 120 clases de razas de perros.
- ➍ **Congelar las capas previas:** Bloquea las capas pre-entrenadas para preservar los pesos aprendidos.
- ➎ **Data Augmentation:** Aplica técnicas de aumento de datos como rotación, volteo horizontal, cambio de brillo, etc., para aumentar la variabilidad del conjunto de datos.

# Plan de Trabajo

- ➊ **Entrenamiento del modelo:** Entrena el modelo utilizando el conjunto de datos Stanford Dogs, incluyendo las imágenes aumentadas.
- ➋ **Fine Tunning (opcional):** Si es necesario, fine tuning desbloqueando algunas capas y reduciendo la tasa de aprendizaje.
- ➌ **Evaluación del modelo:** Evalúa el rendimiento utilizando un conjunto de datos de prueba y métricas como accuracy, precisión, recall y F1-score.
- ➍ **Deployment del modelo:** Hacer Deploy del modelo en Firebase con tensorflow lite y conectarlo a un frontend en flutter para testear.

# Resultados Esperados

- Obtener un accuracy entre el 90% - 95% usando ambos modelos.
- Crear un caso de uso útil para la aplicación de los modelos a la contribución de la disminución de pérdidas de animales domésticos.
- Con lo anterior generar la posibilidad de crear un sistema que permite aumentar el dataset con una clasificación hecha por los usuarios.
- Estudiar la posibilidad de implementar un algoritmo multiclase.