BIG CATS CLASSIFICATION

OVERVIEW

This model aims to accurately classify images of big cats, specifically cheetahs, jaguars, and leopards, based on their visual characteristics. This is using a computer vision system which is capable of distinguishing between these species despite their strong morphological similarities.

METHODOLOGY

Data for this project was sourced using the iNaturalistAPI, which enabled direct retrieval of wildlife observations for cheetahs, jaguars, and leopards. Many of the images were 'research-grade' quality, meaning that there was consensus between ≥ 2 individuals on the image's species. Each image was automatically organised into a directory structure compatible with PyTorch's `ImageFolder` format and subsequently split into training (70%), validation (15%), and testing (15%) subsets.

Initially, a broad range of data augmentations was implemented to promote generalisation, including random crops, colour jittering, MixUp, and CutMix. However, early experiments revealed that these transformations often distorted the fine visual details (e.g. rosette structures and spot density) that are essential for distinguishing between big cats. As a result, the augmentation strategy was progressively simplified to focus only on random resized cropping, horizontal flipping, and normalization. This refinement led to substantial improvement in validation performance and overall stability.

MODEL DETAILS

The final model was built using ResNet-18, a convolutional neural network (CNN) architecture known for its balance between performance and computational efficiency. The network consists of 18 layers, with the final fully connected layer being modified to output 3 classes: cheetah, jaguar, and leopard.

To comply with ResNet-18's configuration, input images were resized and centre-cropped to 224×224. Training was conducted using the Adam optimizer with a cosine annealing learning rate to gradually reduce learning rates and stabilise convergence. Exponential Moving Average (EMA) weight tracking was incorporated to improve model generalization.

Class weighting was applied to the training loss function to correct an imbalance in class predictions. Earlier versions of the model under-predicted cheetahs, leading to poor recall for that class. To address this, the loss weights were set to 1.3, 1.1, and 1.25 for cheetah, jaguar, and leopard respectively. This configuration increased the penalty for misclassifying cheetahs, improving their recall while maintaining balanced performance across all classes.

EVALUATION

Below are the performance metrics used to calculate the efficacy of the model during validation and testing.

Macro-F1
- Model's precision and recall when predicting

RESULTS

Below are the results from testing.

Cheetah	Jaguar	Leopard	Overall
Precision: 0.83	Precision: 0.83	Precision: 0.76	Precision: 0.81
Recall: 0.89	Recall: 0.82	Recall: 0.72	Recall: 0.81
F1-score: 0.86	F1-score: 0.83	F1-score: 0.74	Macro-F1: 0.81