

Project Proposal

Title: Image Classification of the Pet Breeds Using the Oxford-IIIT Pet Dataset

Project Description:

In this project I aim to build a model to classify the images of different pet breeds which contain dogs and cats. I am using Oxford-IIIT Pet Dataset for this task. I will use CNN to learn visual features and distinguish between multiple classes that dataset contains. The project will include preprocessing, model training, hyperparameter tuning, evaluation, finetuning and final model testing on unseen data.

Why is it good:

This dataset is widely used for commercial and research purposes. The dataset is manageable and balanced. It has overall 37 pet's classes. So, I can use subsets of classes for experimentation. Also, the dataset has real world diversity in terms of posing, lightning and background so it will make the task meaningful. Additionally, I can apply the concepts we have learned in class on this dataset and make it a part of my portfolio.

How will I do it:

For this task I plan to do it in the following steps:

- Loading and preprocessing the dataset
- Will be using data augmentation to improve generalization, so model don't overfit.
- Will implement a CNN from scratch as a baseline, followed by pretrained model like ResNet.
- Using techniques like early stopping so models don't overfit.
- Hyperparameter tuning like learning rate batch size
- Evaluate the performance based on metrics like accuracy, precision, recall and visualize using confusion matrix.

What will I use:

For this project I am going to use Oxford-IIIT Pet Dataset which is available on **Datasets — Torchvision 0.17**. The dataset consists of 7000 labeled images which have 37 breeds of pets. I will be using 10 classes to simplify the task however also maintaining the diversity.

How will I evaluate the performance:

First, I will split the data into training, validation and test. The model's performance will be measured through the classification accuracy across on validation and test dataset. Additionally, I will use metric like precision, recall and f1 score which will provide the robustness of the model across different classes. Further confusion matrix and sample visualizations of misclassification and correct prediction will help me to understand the model behavior and areas for improvement.