Deep and Reinforcement Learning - CAP 5619

Term Report

Breed Recognition for Dogs

Team Members:

Avaneeshakrishna Shastry Chakracodi (AC23BI)

Akhil Gorthi Bala Sai (AG23BU)

Abstract:

Dog breed classification using deep learning has become a significant application of computer vision, facilitating tasks such as pet identification and breed-specific care. In this project, we present an approach to dog breed classification utilizing deep learning techniques on the Stanford Dog dataset. With its extensive collection of annotated images spanning over 120 breeds, the dataset offers a diverse and comprehensive resource for training and evaluating deep learning models. Leveraging convolutional neural networks (CNNs) and transfer learning, our method aims to accurately predict dog breeds from images with high precision and recall. We explore various data augmentation techniques and model optimization strategies to enhance performance and generalization capabilities. Through rigorous experimentation, we demonstrate the effectiveness of our approach in addressing the complexities of breed classification while emphasizing the broader implications for animal welfare and pet-related services.

Introduction:

In recent years, strides in deep learning have transformed computer vision, facilitating intricate image recognition tasks that were once considered daunting. Among these tasks is dog breed classification, an endeavour cantered on sorting dog images into distinct breed categories. Beyond mere curiosity for pet owners, this classification holds profound implications in fields like veterinary medicine, animal welfare, and breed-specific studies. By harnessing the meticulously annotated Stanford Dog dataset, this project endeavours to construct a robust deep learning model. The aim is to adeptly discern and classify dog breeds from images, demonstrating the power of contemporary machine learning techniques in addressing real-world challenges.

Dataset:

For our project, we have chosen to utilize the Stanford Dogs Dataset available on Kaggle. This dataset comprises images representing 120 different breeds of dogs sourced from various regions worldwide. It is a curated collection meticulously assembled using images and annotations from ImageNet, specifically tailored for fine-grained image categorization tasks.

- Number of categories: 120
- Number of Images: 20,580
- Annotations: Class labels, Bounding boxes

The contents of the Stanford Dogs dataset include:

Image Variety: The dataset encompasses a diverse range of images, showcasing dogs of different breeds, ages, and poses. This diversity ensures comprehensive coverage of the various characteristics and nuances associated with each breed.

Annotation Information: Each image in the dataset is accompanied by detailed annotations, providing valuable metadata such as breed labels. These annotations play a crucial role in training and evaluating deep learning models for dog breed classification tasks.

Fine-Grained Categorization: The dataset is structured to support fine-grained categorization, enabling the differentiation of breeds that may share similar physical features. This granularity enhances the precision and accuracy of classification models trained on the dataset.

By leveraging the rich and extensive content of the Stanford Dogs dataset, we aim to develop a robust deep learning model capable of accurately identifying and categorizing dog breeds from images. This dataset serves as the cornerstone of our project, facilitating comprehensive exploration and analysis of dog breed classification techniques in the realm of computer vision.

Objective:

Objective of this project is to build a CNN-based deep learning model to categorise dog breeds.

Methodology:

Using ResNet50 Model:

Data Collection and Preprocessing:

From the collected dataset available on Kaggle, we processed first 50 different dog breeds. This is due to the available computational power at hand. This dataset provides a diverse range of dog images, ensuring robust model training. Each image was resized to 224x224 pixels and normalized to match the input requirements of the pre-trained ResNet50 model. We iterate over the first 50 breed directories in the dataset, assigning each a unique numerical label. For each breed, it loads and preprocesses the images, converting them to RGB format and applying transformations like resizing and normalization. The processed images are appended to a list, along with their corresponding numerical labels, facilitating supervised learning. This approach ensures data consistency and prepares the images for input into the deep learning model. Overall, the code efficiently organizes and preprocesses the image data, laying the groundwork for subsequent model training and evaluation. We split the dataset into training and testing sets, with 80% of the data used for training and 20% for testing.

Model Architecture:

ResNet50 is a convolutional neural network architecture renowned for its deep structure, featuring 50 layers. Introduced by Microsoft Research, it leverages residual learning to address the vanishing gradient problem in very deep networks. ResNet50 employs skip connections or shortcuts to skip layers, allowing gradients to flow more directly during training. These residual connections enable the training of deeper networks without encountering degradation issues. Its architecture consists of multiple residual blocks, each containing convolutional layers with batch normalization and ReLU activations. ResNet50 has been widely adopted in various computer vision tasks, demonstrating exceptional performance in image classification, object detection, and image segmentation.

We employed a transfer learning approach, utilizing the pre-trained ResNet50 model available in PyTorch's torchvision library. The ResNet50 model has shown remarkable performance in various image classification tasks due to its deep architecture and skip connections. We modified the model's fully connected layer to adapt it to our dataset, replacing it with a custom fully connected layer to match the number of dog breeds in our dataset. Additionally, we added a dropout layer with a dropout rate of 0.5 to reduce overfitting during training.

Training:

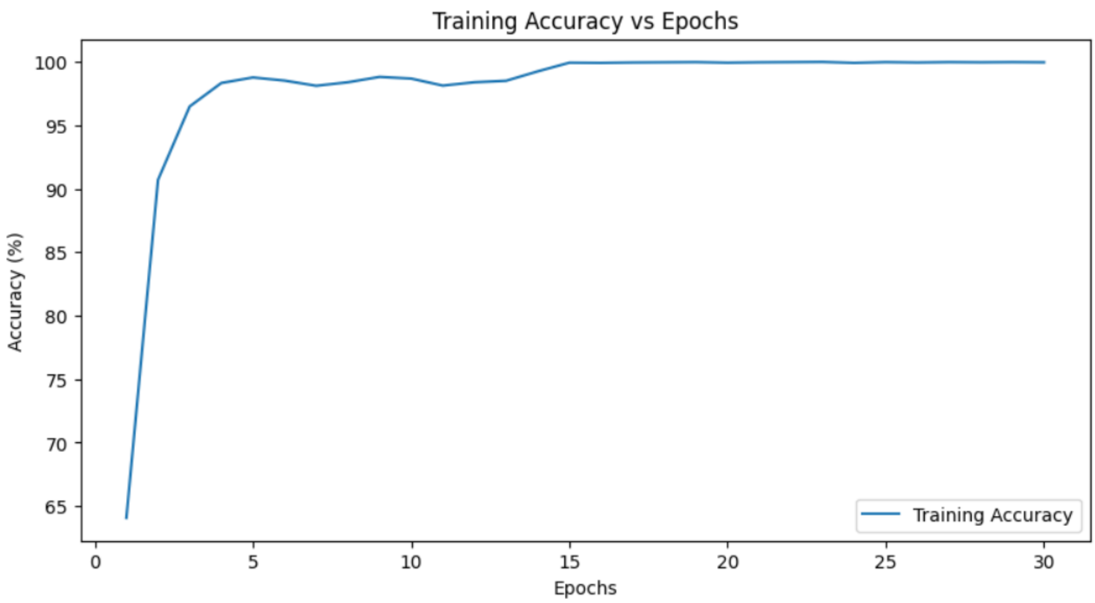
We trained the modified ResNet50 model using the training dataset. The model was trained for 30 epochs using mini-batch stochastic gradient descent. We utilized the CrossEntropyLoss function as the loss criterion and the Adam optimizer with a learning rate of 0.0001. To prevent overfitting and improve generalization, we employed a learning rate scheduler that reduces the learning rate when the validation loss plateaus. During training, we monitored the training loss and accuracy to track the model's performance over epochs.

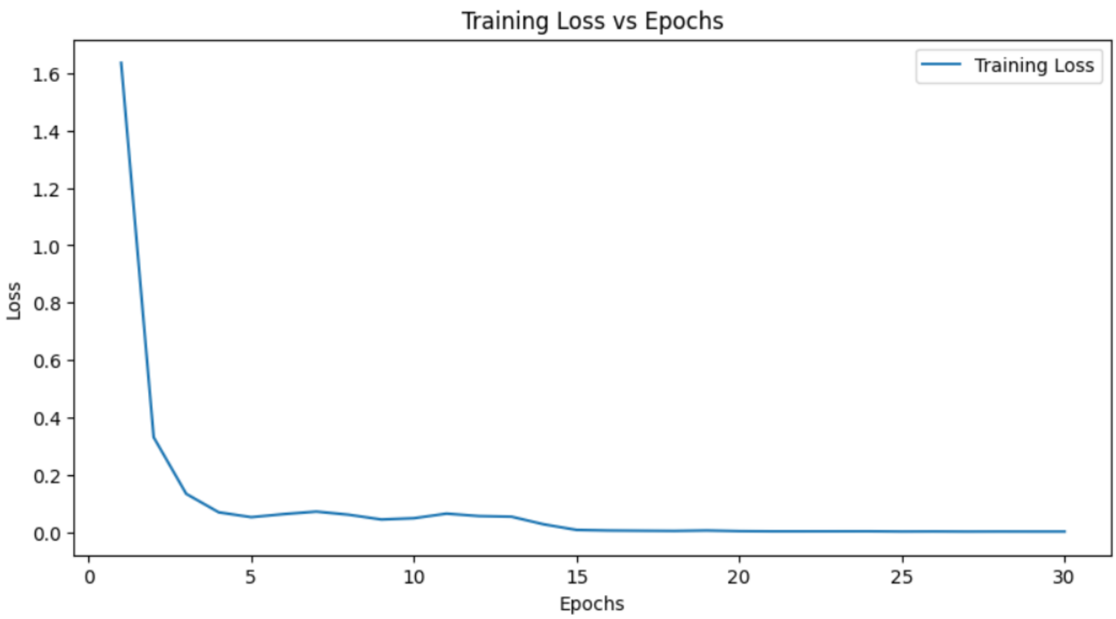
Evaluation:

After training, we evaluated the trained model's performance on the testing dataset. We calculated the test loss and accuracy to assess the model's generalization ability on unseen data. Additionally, we randomly selected images from the testing set to visualize the model's predictions and compare them with the ground truth labels.

Experimental Results:

Using ResNet50 Model:






```
Testing - Loss: 0.5439794140833395, Accuracy: 87.02734147760326 %
Correct Label: n02102040-English_springer
Predicted Label: n02102040-English_springer
```

Example Output:

Conclusion:

The project involves building a deep learning model for breed classification using the Stanford Dogs Dataset. The dataset consists of images of various dog breeds, and the task is to classify these images into their respective breeds. The project utilizes PyTorch and torchvision for model development and evaluation. We have trained and tested our model for first 50 dog breed classes. We were able to achieve train accuracy close to 98% and test accuracy of 87% as shown in the results.

Future Work:

In future, we can include all the classes for training given there is enough computational power at hand. Additionally, increasing the diversity of the dataset by incorporating more dog breeds or augmenting the existing images could improve the model's ability to generalize.

References:

https://www.kaggle.com/datasets/jessicali9530/stanford-dogs-dataset