

Report

Introduction

The main goal of this project is to create a model that will identify the breed of a dog in a picture. The training dataset is a set of pictures labeled by the breed of the dog in the picture. The dataset can be found here:

<https://www.kaggle.com/competitions/dog-breed-identification/data>

Data Preparation

Image data generation was used for data preparation. The following parameters were deemed the most relevant:

```
(rescale=1./255,  
 rotation_range=20,  
 width_shift_range=0.2,  
 height_shift_range=0.2,  
 shear_range=0.2,  
 horizontal_flip=True,  
 fill_mode='nearest',  
 validation_split=0.2)
```

It was decided not to add zooming as there are pictures where zooming would remove the dog from the picture. The size of the input image was set to 400 x 400.

Method

Several architectures (InceptionResNetV2, InceptionV3, VGG16, ResNet50, AlexNet) with an additional dense layer and softmax activation were trained using the Adam optimizer, categorical cross-entropy loss, and a learning rate of 0.001. Among those models, InceptionResNetV2 and InceptionV3 stood out the most. For both models, dropout and global average pooling 2D were added. The best results were achieved by using early stopping based on the validation loss. After hyperparameter fine-tuning, a learning rate of 0.001 and dropout of 0.3 performed the best for both models.

Results

The InceptionV3 model showed a validation loss of 0.3374 with a validation accuracy of 0.9227, while the InceptionResNetV2 model showed a validation loss of 0.2633 with a validation accuracy of 0.8924. Therefore, the InceptionResNetV2 model was used for later predictions.

