

Domain Background

Dog breed classification is a challenging task in ML. This project is on the Image classification using CNN.

If the image of a dog is provided it should identify the dog breed. If image of a human is provided it should provide the most resembling canine breed. This is a multi-class classification problem using supervised machine learning under deep learning CNN.

Problem Statement

As many dogs have similar features with dogs of other breed. It is a challenge to correctly identify the breed. There are 133 dog breeds to be classified

As part of the project two major tasks are:

1. Dog face detection: to detect dog breed when provided with dog image
2. Human face detector: to detect the most resembling dog breed for the provided human image.

Datasets and Inputs

Datasets and input are provided by the Udacity from links

Dog dataset: <https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/dogImages.zip>

Human dataset: <https://s3-us-west-1.amazonaws.com/udacity-aind/dog-project/lfw.zip>

There are total 13233 human images and 8351 dog images. Images are of size 250*250.

Solution Statement

To detect human faces, we use Open CV's Haar feature based cascade classifiers model.

OpenCV already contains many pre-trained classifiers for face, eyes, smile etc. Those XML files are stored in opencv/data/haarcascades/ folder. Using this facedetector we can detect human faces.

To detect dogs, we use a pretrained VGG-16 model. **VGG16** is a 16 layers deep convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition". The model achieves 92.7% top-5 test accuracy in ImageNet.

We then create a CNN from scratch to classify Dog breeds. We should get an accuracy of atleast 10%.

Now create a CNN using transfer learning using resnet18. We should get an accuracy of atleast 60% with this approach.

For transfer learning we use ResNet-18. ResNet-18 is a convolutional neural network that is 18 layers deep. You can load a pretrained version of the network trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories, such as

keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224

We use this model for transfer learning and adapt it to our requirements.

Now we create an algorithm to detect if image is of a dog then return the predicted breed.

If image is of a human then return the resembling dog breed. If image is of neither dog nor human then provide output that indicates an error.

Benchmark Model

The CNN model created from scratch should have atleast 10% accuracy.

The CNN model created using transfer learning should have accuracy of 60%.

Evaluation Metrics

Accuracy will be the main metric used to measure this model.

Project Design

First: Detect human faces using Open CV's Haar feature based cascade classifiers model.

Second: Detect dogs using a pretrained VGG-16 model

Third: Create a CNN to classify Dog breeds from scratch. We should get an accuracy of atleast 10%.

I will use convolution layers, each followed by maxpooling layer. Maxpool layers helps to reduce the dimensions of the feature maps.

The output is then passed to linear layers with dropout layers in between them. Drop out layers help in avoiding overfitting by zeroing few nodes in network.

Fourth: Create a CNN using transfer learning using resnet18. We should get an accuracy of atleast 60%.

The last layer is modified and trained to output only 133 feature outputs. To classify the dog breeds.

Fifth: - Algorithm to detect if image is a dog then return the predicted breed.

- if image is of a human then return the resembling dog breed.

- if image is of neither dog nor human then provide output that indicates an error.