E9 241 Digital Image Processing
Assignment 05 Report

Name: Saptarshi Mandal

S.R. No.: 04-03-01-10-51-23-1-22925

Discipline: Signal Processing

Department: Electrical Engineering

Q1. KNN Classifier with deep features:

Observations/Results:

The Pre-Trained Network chosen for this experiment is VGG16 trained on the 'ImageNet' Dataset which contains 1000 classes and roughly 14 million images.

VGG 16 contains 13 Convolutional Layers and 3 Fully Connected Layers.

For this experiment, the output is received from the last fully connected layer (containing 4096 neurons) of VGG16.

This VGG16 model is used to generate deep features (**vectors of size (1,4096)**) from the given training samples/dataset.

A KNN (**K=3**) classifier is trained on the generated deep features and is tested on the given test samples/dataset.

The accuracy observed on the test dataset is **0.73.**

The confusion matrix (Figure 1) is given below along with a wrong classification (Figure 2):

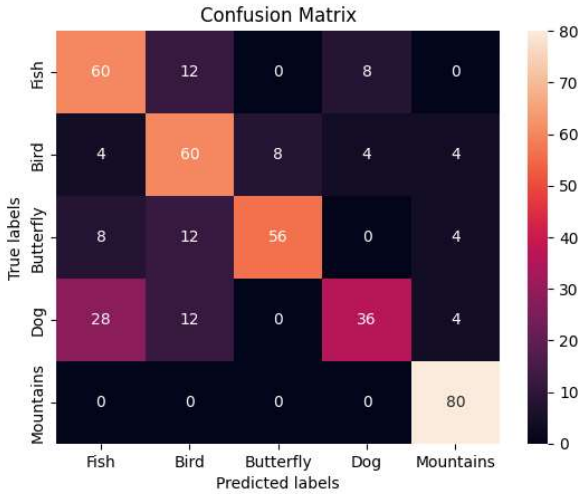

Figure 1

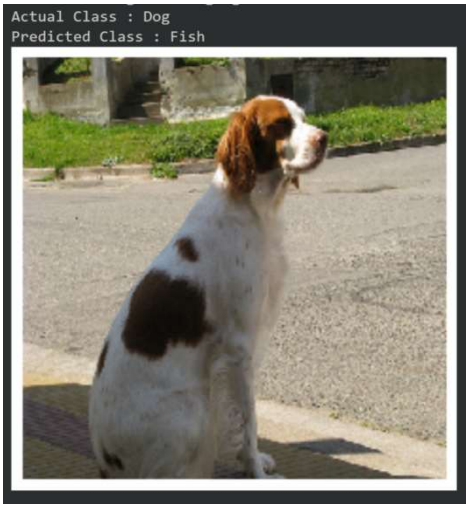

Figure 2

Comments/Inferences:

The VGG16 model is trained on 1000 classes, so its trained to create deep features for best performance with many more classes than the small dataset in this experiment which contains only 5 classes.

Tuning the network on the small dataset might yield better results.

Q2. Fine-tuning:

Observations/Results:

In this experiment, the last fully connected layer is tuned using the training samples of the given small dataset (with 5 classes).

For tuning, the VGG16 model is loaded till the last fully connected layer and all the layers are made 'untrainable'. A fresh new fully untrained fully connected layer is added at the end of the model. Along with this a layer of 5 neurons is also added as the classifying layer.

Now training is done on this model on the given training dataset with a small learning rate (**0.00001**). This tunes the last fully connected layer on our dataset.

Now the output of the last fully connected layer (no. of neurons = 4096) will provide the deep features of any image. Hence, the model is now used to generate deep features from the given training and test dataset.

These deep features are used to train a KNN model (K=3) similar to the previous experiment and are tested on the test dataset.

The accuracy observed on the test dataset is **0.80.**

The confusion matrix (Figure 3) is given below along with a correct classification (Figure 4):

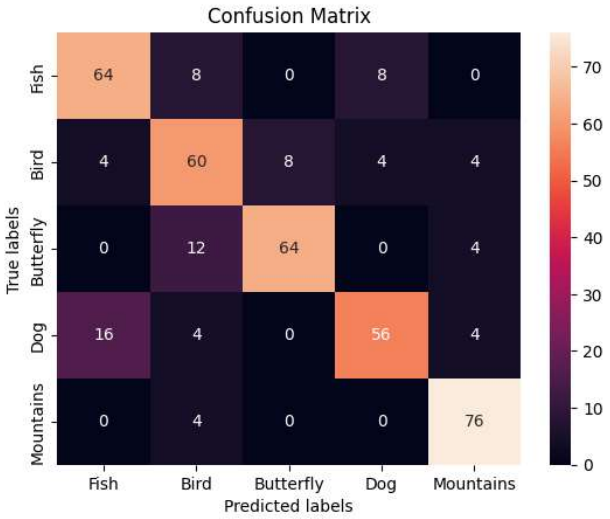

Figure 3

Figure 4

Comments/Inferences:

The VGG16 model is trained on 1000 classes, so it's trained to create deep features for best performance with many more classes than the small dataset in this experiment which contains only 5 classes.

Tuning the network on the small dataset changed the weights and biases of the last fully connected layer which resulted in a better generalization and as expected **provided better accuracy** on the test dataset.