Dog Breed Identification

CSA301 Deep Learning
Bachelor of Science in Information Technology
(Year IV, Semester I)

RESEARCHER(S)

Galey Wangmo (12190051)
Gyeltshen Wangdi (12190053)
Tshering Dorji (12190093)
Ugyen Wangdi (12190105)

GUIDED BY

Tawmo Tawmo

Gyalpozhing College of Information Technology
Gyalpozhing : Mongar

1 Abstract

Dog Breed identification is essential for many reasons, mostly to understand individual breed conditions, health concerns, interaction behaviors, and natural instincts. This paper presents a solution for identifying dog breeds from facial photographs. The proposed method applies a deep learning based approach in order to recognize their breeds. The current paper presents a fine-grained image recognition problem of identifying dog breeds in given images. The network is trained and evaluated on the Stanford Dogs dataset. The proposed method is evaluated using Convolution Neural Networks with various augmentation settings. As of now, the proposed model achieves an accuracy of 53% on the Stanford dataset with 10 dog breeds and a total of 5000 datasets. The study concludes that the model does not perform very well given the number of dataset and given algorithm. Therefore, the number of dataset should be increased to improve classification performance. However, due to the limited number of dataset available, it is better train the model using techniques such as transfer learning.

Keywords: Deep Learning; Convolution Neural Network; Dog Breed Classification; Image Augmentation; Transfer Learning.

2 Introduction

Dogs are popular pets because they are playful, friendly, loyal, and tend to listen. As of 2019, there are approximately 340 breeds recognized by the Federation Cynologique Internationale (FCI)[1], the World Federation of Dog Breeds, known as the World Dog Breeds Organization. However, the American Kennel Club[2] currently only recognizes 192 breeds. Knowing dog's breed is not trivial, as different breeds were evolved to perform different tasks. Some breeds, for instance, were bred to hunt, kill and more likely to bite. Unfortunately, humans have hard time telling them apart, until the dog comes from one of the few popular and distinctive breeds such as golden retriever.

It is possible to identify Dog breeds using an expert-based approach and DNA approaches. However, dog experts with variety of knowledge on different breeds of dogs are not readily available. Besides, the result they provide is not much accurate. Alternatively, DNA test provides precise and accurate result, yet DNA approach is an expensive and time consuming approach.

Now, with the help of neural networks in deep learning, a wide range of problems can be solved. There is virtually no limit to the areas in which this technique can be applied. Some of the popular applications of neural networks today include image/pattern recognition, autonomous vehicle trajectory prediction, and face recognition.

Therefore, the current paper will use a deep learning approach to identify dog breeds. The current paper presents the methodology of fine-tuning CNN which is implemented in Stanford dog breed dataset. The convolution neural network is similar to the deep neural network which has weights and biases. CNN has filters which predicts the specific features or patterns present in the original data. On the whole, this paper will predict the breed of a dog using an image of a dog as an input to the algorithm. The paper will then use a convolutional neural network to output a predicted breed.

3 Related Work

This segment provides the previous attempts that are related to the current research.

Kosin Chamnongthai et al.[3] solves similar problem by finding dog breed face using Coarse-to-Fine Concept and PCA.

Middi Venkata Sai Rishita et al.[4] came up with similar solution to find out the dog breed as well as the resembling dog breed of the human if supplied using CNN.

Similar to the current paper Richard O. Sinnott et al. [5] provide IOS based mobile application for dog breed classification which makes use of CNN and utilizes the big data processing infrastructure.

4 Methodology

4.1 System Overview

The process of building the model to classify dog breeds will start with the collection of data, which will be used from kaggle. After which the dataset will be processed with the necessary features and conversion of the images into pixel data stored into csv files. Then the paper will set up a structure for the dataset, into the training and testing sets.

The model will then be trained with the training set and checked for accuracy. If the model obtain high accuracy(greater than 90%), the model will be tested with the testing set. If the model has low accuracy(less than 90%) then the model will be trained again after tuning the hyper-parameters until the desired accuracy is obtained. The model will be checked for under-fitting and over-fitting and then the model will be improved accordingly.

Lastly, after the model is completed, it will be integrated with the website to classify the dog breeds. The model should be able to classify the unseen dog images accurately.

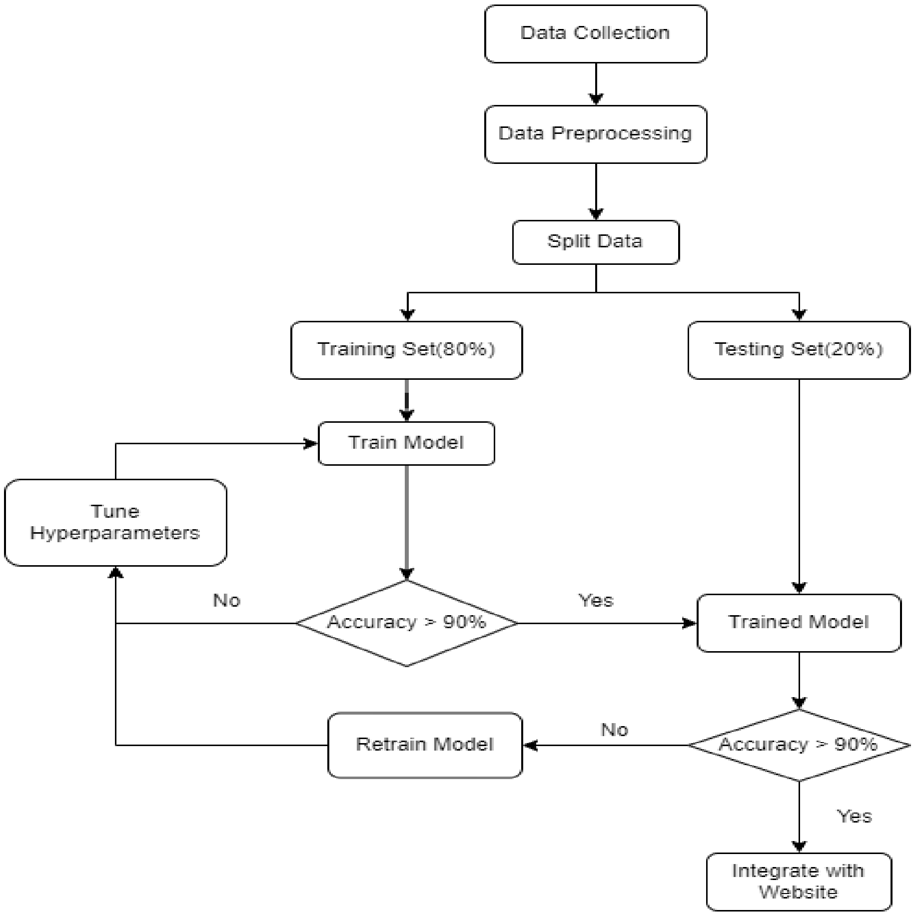

Figure 1: Flowchart

4.2 Algorithm

Convolutional Neural Network

A CNN architecture-based approach is proposed for Dog Breed Identification. A three convolutional layers CNN to classify dog breeds is shown in Fig.2. Similarly, the model will consist of several convolutional layers, pooling layers and fully connected layers between the input and output layers based on how the model performs while training.

CNNs[6] are particularly useful for image classification, image recognition, and computer vision (CV) applications, as they provide highly accurate results, especially when processing large amounts of data. Object characteristics are also absorbed over time as the object data passes through the many layers of the CNN. This direct learning eliminates the need for manual feature extraction. The steps involved in image classification are an input layer, a hidden layer, and an output layer.

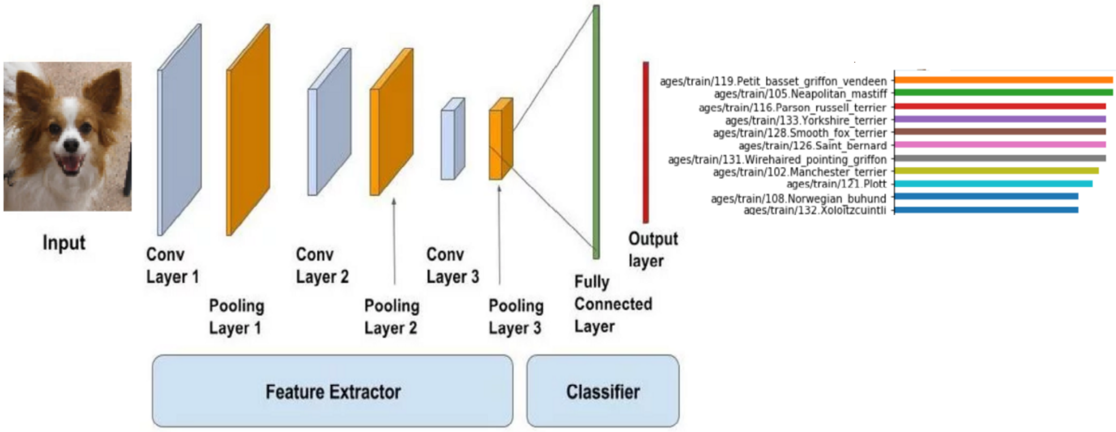

Figure 2: Sample CNN architecture

Some of the parameters that is used are as follows:

1. Padding
   Padding is a term relevant to convolutional neural networks as it refers to the amount of pixels added to an image when it is being processed by the kernel of a CNN. We are going to set our padding to zero as that allows the size of the input to be adjusted to our requirement. Also, it is mostly used in designing the CNN layers when the dimensions of the input volume need to be preserved in the output volume.

2. Stride
   Stride is a parameter of a neural network's filter that modifies the amount of movement over the image or video. For example, if a neural network stride is set to 1, the filter will move 1 pixel, or unit, at a time. We will use one stride as we are going to preserve the spatial size with padding.

3. Filter
   Filters -also known as kernels- scan the image region after another. The region size is determined by the window or kernel size. We are going to use 3X3 kernel size as the overall input size would be much efficient when the kernel size is small. Thereby consuming less time to process and there would be less ambiguities.

4. Batch size
   A batch size in a neural network is a hyperparameter that defines the number of samples to work through before updating the internal model parameters. It is the number of samples that will be passed through to the network at one time. We will be taking 32 as our batch size as it is the optimal batch size.

5. Epochs
   An epoch means training the neural network with all the training data for one cycle. In an epoch, we use all of the data exactly once. A forward pass and a backward pass together are counted as one pass: An epoch is made up of one or more batches, where we use a part of the dataset to train the neural network.

4.3 Dataset

For this project, the group will use an existing dataset of dog breeds from kaggle[7] and will be classifying a total of 10 different classes of dog breeds. As for the image dimension, 256 * 256 * 3 will be used as an input. The channel will be taken as 3 as the model will not be using grey-scale, since different classes of dog breeds have different colors.

The dataset contains a total of 5000 images:

1. 10 classes of dog breeds.
2. 500 images each will be split over training (80%) and testing (20%) sets.

4.4 Evaluation Metrics

Model accuracy in deep learning technologies can be evaluated using a confusion matrix. It is used to describe a model's performance based on test data where true values are known. It is also used to compute recall, precision, accuracy, and f1 scores from the confusion matrix using:

1. True Positive (TP): when the trained model correctly predicts the positive class.
2. True Negative (TN): when the model correctly predicts the negative class.
3. False Positive (FP): when the model incorrectly predicts the positive class.
4. False Negative (FN): when the model incorrectly predicts the negative class.

The Accuracy is ratio of correctly predicted class to the total instances. This way of finding accuracy is best known when there is an equal number of instances in each class in the datasets. The given equation is used to find the accuracy:

$$Accuracy = \frac{TP}{TP + TN + FN + FP}$$

The Precision is used to compare True Positive(TP) and False Positive(FP) entities, it is calculated from the confusion matrix using following formula:

$$Precision = \frac{TP}{TP + FP}$$

The F1-Score is a solution when there is a point where output estimation with recall and precision is no longer possible. It takes the precision and recall values and averages them, the formula is:

$$F1 - Score = \frac{TP}{TP + FN}$$

TP are for entities that are correctly categorized, while FP are the entities that are incorrectly classified. The Recall is used to compare the TP entities to FP entities that are not labelled at all, it is given by the formula:

$$Recall = \frac{2xPrecisionxRecall}{Precision + Recall}$$

5 Results and Discussions

This section of the paper presents the results of the model for the first draft. Paper will be discussing the evaluation in terms of accuracy, precision, recall, f1-score, and confusion matrix. Accuracy is monitored on the training and test datasets; this metric represents the mean percentage of correctly classified classes on a dataset.

After selecting the best loss and optimizer for the model, the model is then trained using that loss function and optimizer. The model is then evaluated to measure the performance or the accuracy on the training and testing dataset using the score and losses. Following plots of loss and accuracy on the training and validation sets were being plotted. The plot helped to check the existence of 'overfitting' and 'underfitting' by plotting the differences between training and validation accurac. Besides, it helped monitor if there is an improvement in the training accuracy.

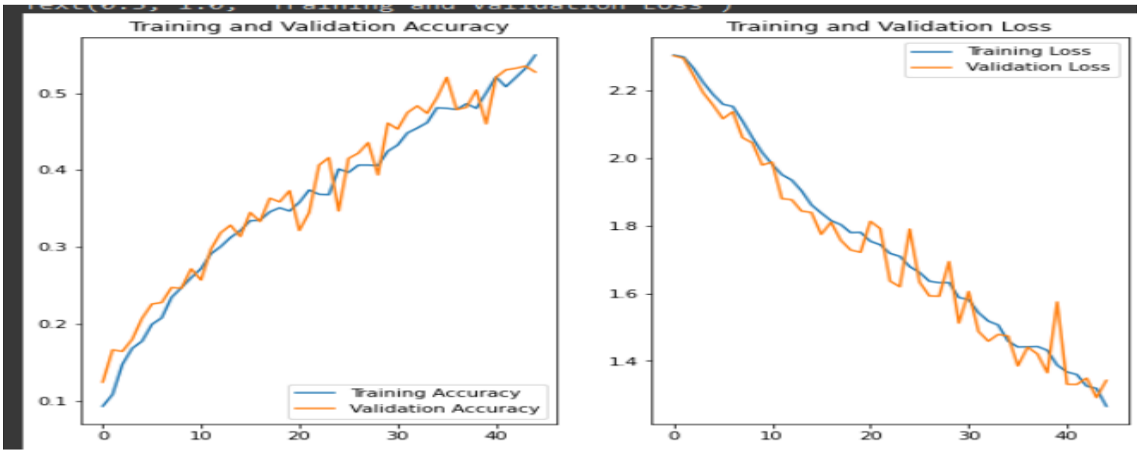

Figure 3: Accuracy and loss graph

The current dog breed model achieved 55.68% for training and 52.73% accuracy on the test dataset. The following figure is the classification report generated from predicting the test dataset which shows the precision, recall and f1-score, macro avg, weighted avg and the overall accuracy (53%).

	precision	recall	f1-score	support
0.0	0.57	0.56	0.57	108
1.0	0.42	0.46	0.44	93
2.0	0.51	0.52	0.52	98
3.0	0.56	0.63	0.60	101
4.0	0.57	0.51	0.54	110
5.0	0.66	0.51	0.58	96
6.0	0.37	0.53	0.43	95
7.0	0.48	0.37	0.42	126
8.0	0.53	0.56	0.55	101
9.0	0.70	0.65	0.67	96
accuracy			0.53	1024
macro avg	0.54	0.53	0.53	1024
weighted avg	0.54	0.53	0.53	1024

Figure 4: Accuracy, Macro avg, weighted avg

Evaluating the model using confusion matrix: The paper includes evaluation of the model by plotting a confusion matrix for predicted labels against the actual label of test images. To visualize the number of images that got predicted correctly and the number which were mispredicted, here the confusion matrix from matplotlib is being used. The matrix helps to see the number of test samples which were misclassified. From the matrix below, the maximum of images got predicted correctly although the model cannot be considered good.

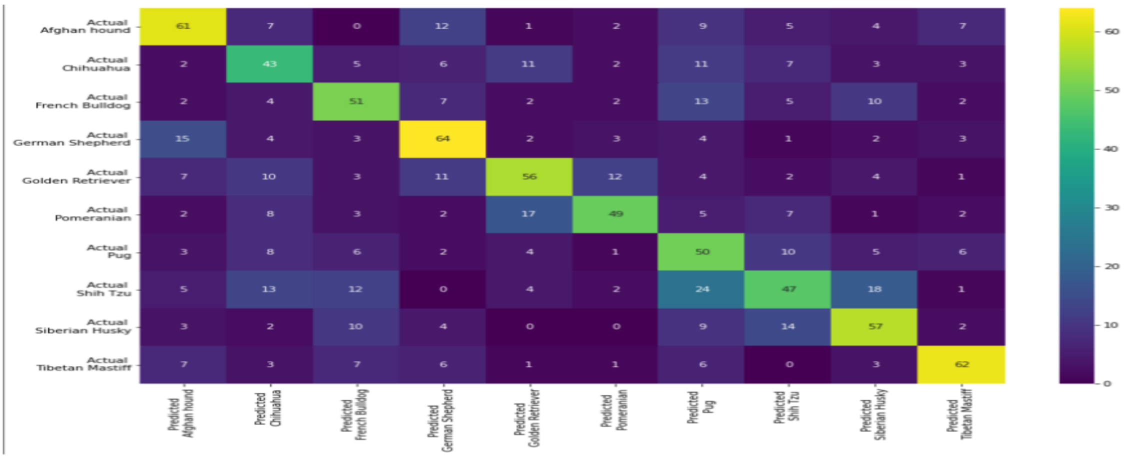

Figure 5: Confusion matrix

The figure given below shows the predictions made by the model with the actual class, predicted class and the confidence score of the model.

Figure 6: Prediction result

The figure given below shows the model after being integrated with the Dog Breed Identification website.

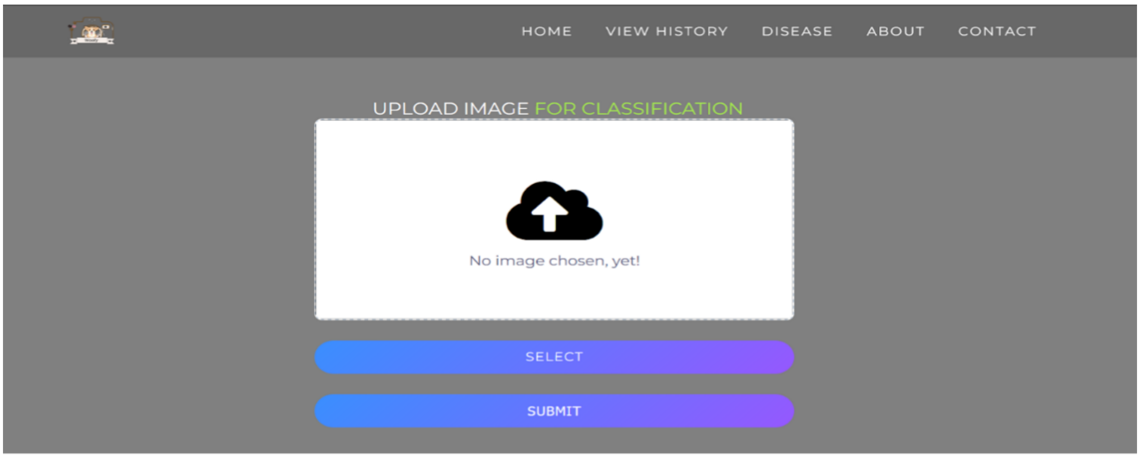

Figure 7: Home Page

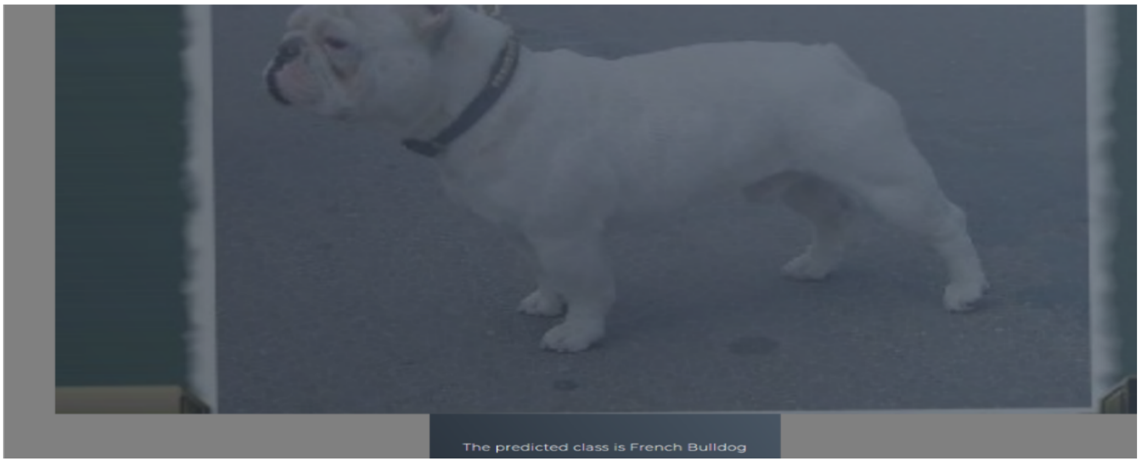

Figure 8: Predicted result

6 Conclusion

Lastly, from the results and evaluations of the model, 'SparseCategoricalCrossentropy' loss function gave the higher accuracy and minimum loss for the CNN model. In order to produce and deploy the final model, it will again be trained by adjusting some hyper-parameters such as batch sizes, number of epochs, learning rates, adding dropout layers, and batch normalizations until best accuracies and minimum losses are achieved. Further, the number of samples for each class will be increased and the model will be fine-tuned using techniques such as transfer learning.

References

[1] S. Paatsama and J. Morgan, “Fédération cynologique internationale (fci),” *Hereditary Bone and Joint Diseases in the Dog-Osteochondroses, hip dysplasia, elbow dysplasia-. Hannover: Schlütersche*, pp. 263–266, 2000.

[2] R. J. Huss, “Pups, paperwork, and process: Confusion and conflict regarding service- and assistance animals under federal law,” *Nevada Law Journal*, vol. 20, no. 2, p. 11, 2020.

[3] M. Chanvichitkul, P. Kumhom, and K. Chamnongthai, “Face recognition based dog breed classification using coarse-to-fine concept and pca,” in *2007 Asia-Pacific Conference on Communications*. IEEE, 2007, pp. 25–29.

[4] M. V. S. Rishita and T. A. Harris, “Dog breed classifier using convolutional neural networks,” in *2018 International Conference on Networking, Embedded and Wireless Systems (ICNEWS)*. IEEE, 2018, pp. 1–7.

[5] R. O. Sinnott, F. Wu, and W. Chen, “A mobile application for dog breed detection and recognition based on deep learning,” in *2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT)*. IEEE, 2018, pp. 87–96.

[6] V. Veeramsetty, G. Singal, and T. Badal, “Coinnet: platform independent application to recognize indian currency notes using deep learning techniques,” *Multimedia Tools and Applications*, vol. 79, no. 31, pp. 22 569–22 594, 2020.

[7] A. Khosla, N. Jayadevaprakash, B. Yao, and L. Fei-Fei, “Stanford dataset for dog breeds identification.”