Dog Breed Detection

K. Sai Teja Reddy
(B.E) Student
Computer Science
Engineering, at
Birla Institute of Technology and Science
Dubai
f20190036@dubai.bits-pilani.ac.in

Nipun Siddineni
(B.E) Student
Computer Science
Engineering, at
Birla Institute of Technology and Science
Dubai
f20190047@dubai.bits-pilani.ac.in

Siddharth Bhandary
(B.E) Student
Computer Science
Engineering, at
Birla Institute of Technology and Science
Dubai
f20190218@dubai.bits-pilani.ac.in

Melvin Jose
(B.E) Student
Biotechnology
Engineering, at
Birla Institute of Technology and Science
Dubai
f20190121@dubai.bits-pilani.ac.in

***Abstract*— This report consists of an Introduction which explains the usefulness and need for a Dog Breed Detection System, Data Collection which explains the use of ImageNet dataset for the project and how the model is pretrained, Theoretical Background which elaborates on the workings of the ResNet50V2 which is the Convolutional Neural Network used for this project, Results & Discussion which elaborates further on the input initialization, ResNet50V2, augmentation and the results of the project wherein the model gets an accuracy of 80%. Finally the report consists of a Conclusion which summarizes the results of the project, the libraries used & the usefulness of ResNet50V2. The end of the report consists of a list of reports that have been used for Reference.**

Keywords*— *CNN, Convolutional Neural Networks, ResNet, Residual Network, Image Processing, TensorFlow, Fuzzy Logic, Neural networks, ImageNet, ResNet50V2, Image Augmentation, Training

I. Introduction

Neural networks are known to be a series of algorithms which strive to identify relationships between different sets of data in the same way as how a human brain works to identify and correlate different information that is presented to it.

Like the human brain, the neural network consists of neurons, which in this case is a mathematical function which stores and transfers information according to the architecture of the neural network.

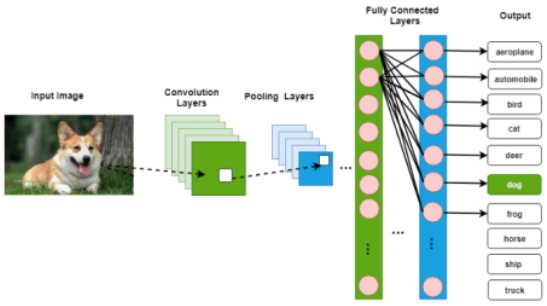

Figure 1 - Simplified representation of CNN

In this project our primary goal is the classification of dog breeds based on their appearance and various physical traits that are prevalent. The importance of creating a neural network model for classifying dog breeds is that it allows dog owners and possible buyers to determine which dog breed is financially feasible for them and how to take care of it, since various dog breeds are known to have different diets and grooming requirements for their fur and general hygiene.

Dog Breed classifications can also aid owners to find dogs that they have lost by easily identifying their dogs based on their unique physical features and

breed. Another instance where Dog Breed classification can also allow corporations to identify which dog breeds are known to be popular to the general population. This can lead these corporations to create targeted advertisements which feature popular dog breeds to obtain a potential consumer for their product.

In the case of veterinarian doctors dog breed classifications can help aid these professionals in studying the physiology as well as the anatomy of different dog breeds which can help aid in these animals' treatment. It also helps in identifying various other dog breeds that can be used for drug testing so as to produce the intended results of the medicine.

In this project we use a pre-trained network known as ResNet (Residual Network). ResNet is known to use pooling layers, convolutional layers, activation layers and finally fully connected layers stacked with each other.

II. Data Collection

ResNet is usually first pre-trained on a dataset of at least million images. In our case we use ten thousand images for test and training data each. For pre-training a model, there are certain steps which are followed -

- Data Preparation - The images are first prepared properly by formatting them and scaling the images to a proper size and then sorting the images into training set, validation set and testing set.
- Initialization - Weights are given randomly when initializing the model first where the randomly given weights are calculated using a normal distribution curve with a mean of 0 and a standard deviation of 0.1.
- Training - The model is trained using mini-batch gradient descent which is a type of stochastic gradient descent (SGD). Here the training set is segregated into smaller batches and then the model is trained with each batch. After each iteration of batch training, the weight gets updated.
- Regularization - To prevent overfitting, regularization approaches such as drop out, weight decay and data augmentation
- Optimization - This is done by changing some parameters in the model such as learning rate, batch size and number of epochs.
- Evaluation - The effectiveness of the model after pre-training is completed is evaluated using the validation data. In this, the model's accuracy, precision, recall and F1 score is evaluated.

After the model has been pre-trained, it can be further trained on a specific dataset after fine-tuning the model to improve the performance by changing weights based on the new dataset. In this case, the ResNet model used has been pre-trained on the ImageNet dataset.

ImageNet dataset was created in 2009 by a set of researchers from Stanford University and Princeton University. It has been used as the gold standard dataset to train models used for image classification. The dataset contains around 14 million photographs which cover around 1000 categories like animals, objects, sceneries etc.,. Each image in the dataset is tagged with an annotation, which is a label describing the image. The images are supposedly a representation of everyday objects and environment.

The dataset is divided into two subsets which are test set and train set. The images in the test set varies massively from the images prevalent in the train set and these images are used to calculate the performance of the model on which the dataset is used to pre-train.

ImageNet dataset has played an essential role in the developing image classification and deep learning models. Some of the most effective models which have been trained/pre-trained with ImageNet dataset are - AlexNet, VGG, ResNet, Inception etc.,.

This model after being pre-trained on the ImageNet dataset has been further trained on the Stanford Dogs dataset. This dataset was developed by researchers at Stanford University in 2011 as part of the Large Scale Visual Recognition Challenge.

The dataset used contains a total of 120 unique images of dog breeds with an average of atleast 160 images per dog breed in the dataset. Each image is labelled with the breed of the dog as well as a bounding box which indicates where the dog is in the image. The breeds included in the dataset vary from having the most common breeds such as Retrievers, Poodles, German Shepherds to having the less common breeds such as Kuvacz, Azawakh etc.,.

	A	B
4	001cdf01b096e06d78e9e5112d419397	pekinese
5	00214f311d5d2247d5dfe4fe24b2303d	bluetick
6	0021f9ceb3235effd7fcde7f7538ed62	golden_retriever
7	002211c81b498ef88e1b40b9abf84e1d	bedlington_terrier
8	00290d3e1fdd27226ba27a8ce248ce85	bedlington_terrier
9	002a283a315af96eaea0e28e7163b21b	borzoi
10	003df8b8a8b05244b1d920bb6cf451f9	basenji
11	0042188c895a2f14ef64a918ed9c7b64	scottish_deerhound
12	004396df1acd0f1247b740ca2b14616e	shetland_sheepdog
13	0067dc3eab0b3c3ef0439477624d85d6	walker_hound
14	00693b8bc2470375cc744a6391d397ec	maltese_dog
15	006cc3ddb9dc1bd827479569fcdc52dc	bluetick
16	0075dc49dab4024d12fafe67074d8a81	norfolk_terrier
17	00792e341f3c6eb33663e415d0715370	african_hunting_dog
18	007b5a16db9d9ff9d7ad39982703e429	wire-haired_fox_terrier
19	007b8a07882822475a4ce6581e70b1f8	redbone
20	007ff9a78eba2aebb558afea3a51c469	lakeland_terrier
21	008887054b18ba3c7601792b6a453cc3	boxer
22	008b1271ed1addaccf93783b39deab45	doberman
23	008ba178d6dfc1a583617470d19c1673	otterhound
24	009509be3ca7cce0ff9e37c8b09b1125	otterhound
25	0097c6242c6f3071762d9f85c3ef1b2f	bedlington_terrier
26	00a338a92e4e7bf543340dc849230e75	dingo
27	00a366d4b4a9bbb6c8a63126697b7656	golden_retriever
28	00a862390341c5be090dd72bd2bc19ef	standard_schnauzer
29	00b7d114bc5166a629a3cc03d9329120	irish_water_spaniel

Figure 2 - List of breeds included in dataset

This dataset has become a benchmark for testing models designed for very fine grained recognition such as identifying between similar looking objects - like identifying the breed of the dog from a given image.

III. Theoretical Background

The identification of dog breeds is not something new and sometimes it might take time to sort them without the help of computers. Computers can identify from the image data given to them and can lower the time constraint of identifying dog breeds. We can use machine learning techniques like Neural Networks to identify different breeds of dogs swiftly and accurately.

Neural Networks are similar to the neurons of a brain that activate when the input is higher than a certain threshold level of the neuron. Neural networks can have many layers in between them to get the desired output. Nowadays, many tech companies use the Neural Networks model for various purposes.

Identification of dog breeds is done by extracting and identifying different features of the dog like color of the dog, ears of the dog and more. These features are different for different dog breeds. First we train the model to accurately identify what dog breed the image belongs to from the features fed to it. When the results are promising, we can use this model to detect various dog breeds accurately.

For image identification we use Convolution Neural Networks. This is one of the most popular forms of image identification neural networks. There are many techniques to identify dog breeds but in this paper we used ResNet50V2 to detect dog breeds.

ResNet50V2 is usually used for image classifications in Convolution Neural Networks. ResNet50V2 belongs to the ResNet architecture and it has 50 layers and is used to train deep neural networks and allow the training of neural networks. The "V2" in the ResNet50V2 implies an improved and better version than the ResNet architecture.

ResNet is known to solve various problems which are complex in nature as the layers in ResNet are known to be pre trained on a huge quantity of images from the ImageNet database. It uses Batch Normalization and RELU activation before multiplying the input by convolutional processes, will be used in this project. To create the dog breed classifier we use python modules such as python 3.8.5, tensorflow, numpys, pandas, sklearn, opencv and matplotlib.

IV. Results & discussion

In our model, we have used the neural network model ResNet50V2. There are various trained parameters in the imagenet dataset which we will be importing into our program code. In our code, we have we have replaced it with the input at hand, which is 224 x 224 pixels instead of using the output of the pre-trained neural network. This input has a shape of m*n*3 dimensions.

The various libraries used are tensorflow, cv2, numpy, and pandas. There are a total of 120 unique dog breeds in the dataset but we will be restricting to using only the top 60 most populated dog breeds among the given dataset for testing and training purposes.

A. Input initialization

We have transformed the input image which in this case is a Rottweiler into a numerical representation so that the neural network can accept the image. A pixel is the smallest unit of data in a picture and many of these pixels together are used to make images. In colored pictures, the image is composed of the three primary hues "RED," "GREEN," and "BLUE." As a result, colored pictures are created using three matrices or channels.

Figure 3 - Input Image of Rottweiler Dog Breed

Pixels are independent components in these matrices, and the matrix's size is determined by the number of pixels. The pixel value, which runs from 0 to 255, represents the brightness or intensity of a pixel. White is represented by a number closer to 255, whereas black is represented by a number closer to 0. There are multiple formats in which the pixels of the image can be stored such as Grayscale, RGB, HSV, CMYK. In our program, we have used the RGB format.

Using the opencv library, we can resize all the images in our library to be of the 224 * 224 pixel format so that it can be a suitable input for our code. Also, all the values in the matrix array will be in the range of -1 and 1.

B. ResNet50V2

ResNet50v2 is a convolutional neural network (CNN) used for picture categorization. It's an upgraded version of the original ResNet design, which debuted in 2015. The "50" in ResNet50v2 refers to the network's layer count, indicating that it is a deep neural network.

ResNet50v2 employs skip connections to facilitate information flow across the network, hence avoiding the vanishing gradient problem that can arise in deep neural networks. Batch normalization is also used, which normalizes the inputs to each layer to increase training speed and generalization performance.

C. Augmentation

We have also used augmented dog images, that is, create more images artificially during the run time for training and testing purposes, so that we can generalize the neural network and help improve its performance.

There are 2 types of augmentation: position augmentation and colour augmentation.

- In position augmentation, the position of the pixels in the input matrix are shifted using concepts such as translation, scaling, cropping, etc.
- In colour augmentation, the changes in the brightness, the contrast, the hue of the image leads to change in pixel value of the input matrix.

D. Results

We have used a database which has 10,222 images in the training dataset for the constructed neural network at hand and have used a database which has 10,357 images in the testing dataset for the designed neural network. We have segregated these datasets with an 80:20 ratio for testing as well as training respectively. This means that 80% of the data from the training dataset will be used to train the neural network and 20% of the data from the testing dataset will be used to test the neural network and adjust weights accordingly.

A learning rate of 0.03 has been used for this neural network and a batch size of 64 images has been selected for 20 epochs. The accuracy obtained has been recorded to be 80%.

```
Epoch 17/20
64/64 [==============================] - 149s 2s/step - loss: 0.7370 -
accuracy: 0.7745 - val_loss: 0.7247 - val_accuracy: 0.7988
Epoch 18/20
64/64 [==============================] - 149s 2s/step - loss: 0.6743 -
accuracy: 0.7959 - val_loss: 0.7362 - val_accuracy: 0.7871
Epoch 19/20
64/64 [==============================] - 152s 2s/step - loss: 0.6526 -
accuracy: 0.8018 - val_loss: 0.7382 - val_accuracy: 0.8076
Epoch 20/20
64/64 [==============================] - 152s 2s/step - loss: 0.6650 -
accuracy: 0.7998 - val_loss: 0.7593 - val_accuracy: 0.8027
WARNING:absl:Found untraced functions such as _jit_compiled_convolution
_op, _jit_compiled_convolution_op, _jit_compiled_convolution_op, _jit_c
ompiled_convolution_op, _jit_compiled_convolution_op while saving (show
ing 5 of 54). These functions will not be directly callable after loadi
ng.
1/1 [==============================] - 1s 872ms/step
Predicted Breed for this Dog is : rottweiler
PS C:\Users\sudhi\Downloads\dog-breed-identification>
```

Figure 4 - Result of project wherein accuracy is 80.27% and the input image has been identified as rottweiler

V. Conclusion

Classifying the various dog breeds available around us is a time consuming process and recognizing a specific dog breed based on its characteristic features can be a challenging task as well, but with the rise of the computer age, we have access to concepts such as neural networks which have made tasks such as dog breed classification programmable.

Neural networks, through repeated training and testing, can successfully extract and identify various characteristic features of a dog, such as the eye colour, ear shape, face structure, etc.

For a computer or program to identify an image, a Convolutional Neural Network (CNN) is commonly used to perform the task and in this research paper, we have used the ResNet50v2 neural network. The ResNet50v2 architecture has been tested and trained on a vast number of images from the ImageNet database.

This helped in making the neural network to solve more complex problems as well. Python modules such as TensorFlow, NumPy, Pandas, scikit-learn, OpenCV, and Matplotlib provided a great number of resources which were used to create a neural network which can identify various dog breeds with promising accuracy.

The input image was successfully transformed into a matrix which is a numerical representation of the image and this matrix was used as input for the ResNet50v2 neural network. Both forms of augmentation, that is, colour augmentation and position augmentation were also used during the runtime of the code which have helped in increasing the accuracy of the neural network.

The datasets have been separated in an 80:20 ratio where 80% of the images from the training dataset were used to train the images and 20% of the images from the testing dataset were used to test the neural network. In our project a learning rate of 0.03 and a batch size of 64 images have also been implemented for the constructed neural network. The testing dataset has 10,357 images and the training dataset has 10,222 images. After multiple tests, the accuracy of the designed neural network was recorded to be 80%

VI. Contributions

Siddharth Bhandary (2019A7PS0218U) - Abstract, Keywords, Introduction, References

Melvin Jose (2019A9PS0121U) - Data Collection, References

K Sai Teja Reddy (2019A7PS0036U) - Theoretical Background, References

Nipun Siddineni (2019A7PS0047U) - Results and Discussion, Conclusion, References

VII. References

[1] Z. Ráduly, C. Sulyok, Z. Vadászi and A. Zölde, "Dog Breed Identification Using Deep Learning", 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, Serbia, 2018, pp. 000271-000276, doi: 10.1109/SISY.2018.8524715

[2] S.P. Vaidya, PVSSR. Chandra Mouli, and V.R Kishore, "A Novel Dog Breed Identification using Convolutional Neural Network", PriMera Scientific Engineering, 2022, pp. 16-21, doi: 10.56831/PSEN-02-023

[3] P. Borwarnginn, W. Kusakunniran, S. Karnjanapreechakorn and K. Thongkanchorn, "Knowing Your Dog Breed: Identifying a Dog Breed with Deep Learning", International Journal of Automation and Computing, vol. 18, no. 1, 2021, pp. 45-54, doi: 10.1007/s11633-020-1261-0

[4] B. Vijaya Kumar, K. Bhavya. "Dog Breed Identification with Fine tuning of Pre-trained models". International Journal of Recent Technology and Engineering (IJRTE),8(2S11),2019,pp.181-184https://www.citefactor.org/journal/pdf/Dog-Breed-identification-using-CNN-Architecture.pdf

[5] Dr. K. Gupta, S. Goel, B. Aggarwal, R. Gupta. "Dog Breed identification using CNN Architecture".International Journal of Scientific & Engineering Research Volume 11, Issue 6, June-2020 1040ISSN2229-5518 https://www.ijrte.org/wp-content/uploads/papers/v8i2S11/B14640982S1119.pdf

[6] W. Shi, J. Chen, M. Liu, F. Liu, "Dog Breed Identification", International Journal of Recent Technology and Engineering (IJRTE), 2018 http://noiselab.ucsd.edu/ECE228_2018/Reports/Report18.pdf

[7] P. Haswar, A. Iyer, P. Godhane, A. Jadhav, and A. Malhotra , "Dog Breed Identification and Age Detection using Neural Networks," Apr-2022. [Online],2022,https://ijirt.org/master/publishedpaper/IJIRT154419_PAPER.pdf

[8] P. Haswar, A. Iyer, P. Godhane, A. Jadhav, and A. Malhotra , "Dog Breed Identification and Age Detection using Neural Networks," Apr-2022. [Online],2022,https://ijirt.org/master/publishedpaper/IJIRT154419_PAPER.pdf

[9] S. I. Kadari, S. S. Kulkarni, and S. G. Kulkarni, "Dog Breed Prediction using Convolutional Neural Network,"Jun-2020,[Online],2020 https://www.ijeat.org/wp-content/uploads/papers/v9i5/D8058049420.pdf

[10] "Dog Breed Classification Using Part Localization", J. Liu, A. Kanazawa, D. Jacobs, and P. Belhumeur, 2012. Classification of dog breeds using part localization. A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid (eds. 2012's Computer Vision - ECCV. Volume 7572 of Lecture Notes in Computer Science. Berlin, Heidelberg: Springer,2012 https://doi.org/10.1007/978-3-642-33718-5_13

[11]CS231n: Convolutional Neural Networks for Visual Recognition, D. Hsu (2015) "Using Convolutional Neural Networks to Classify Dog Breeds" http://cs231n.stanford.edu/reports/2015/pdfs/fcdh_FinalReport.pdf

[12] AK. Agarwal, V. Kiran, RK. Jindal, D. Chaudhary and RG. Tiwari , "Optimized Transfer Learning for Dog Breed Classification",International Journal of Intelligent Systems and Applications in Engineering 10.1s(2022):2022,pp.18-22 https://www.ijisae.org/index.php/IJISAE/article/download/2233/816

[13] K. Uma1, C. Ramesh Kumar and Thirumurugan Shanmugam, "Dog Breed Classifier Using Non-Linear Convolutional Neural Network (CNN)", YMER || ISSN : 0044-0477 VOLUME 21 : ISSUE 10 (Oct) - 2022, pp.682-694 https://ymerdigital.com/uploads/YMER211095.pdf

[14] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2009.

[15] Li Fei-Fei, Bangpeng Yao, Nityananda Jayadevaprakash and Aditya Khosla. Novel dataset for Fine-Grained Image Categorization. First Workshop on Fine-Grained Visual Categorization (FGVC), IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011