**TITLE OF PBL PROJECT**

**DOG BREED IDENTIFICATION**

**A PBL Project Report**

**Submitted by:**

**Parangat Narsingh Pradhan (2018009122)**

**Rijesh Shrestha (2018011897)**

in partial fulfillment for the award of the degree

of

**BACHELOR OF TECHONOLOGY**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

At



**SHARDA UNIVERSITY**

**Knowledge Park III, Greater Noida, Uttar Pradesh 201310**

**I**

**TABLE OF CONTENTS**

Title Page I

Declaration of the Student II

Certificate of the Guide III

Abstract IV

Acknowledgement V

List of Figures VI

List of Tables VII

**1. INTRODUCTION**

1.1 Problem Definition 1

1.2 Project Overview 2

1.3 Hardware Specification 3

1.4 Software Specification 3

**2. LITERATURE SURVE**

2.1 Basic Definitions1

2.2 Existing System 2

2.3 Proposed System 3

2.4 Feasibility Study 4

**3**. **SYSTEM ANALYSIS & DESIGN**

3.1 Requirement Specification 1

3.2 Flowcharts / DFDs 2

3.3 Design and Test Steps / Criteria 3

3.4 Algorithms and Pseudo Code 4

**4. RESULTS / OUTPUTS** 1

**5. CONCLUSIONS**  2

**6. REFERENCES** 3

**DECLARTION OF STUDENT**

We Parangat Narsingh Pradhan and Rijesh Shrestha solemnly declare that the project report is based on our own work carried out during the course of our study under the supervision of Ms. Kanika Singla (SET Assistant Professor).

We assert the statements made and conclusions drawn are an outcome of our research work. We further certify that

I. The work contained in the report is original and has been done by us under the general supervision of our supervisor.

II. The work has not been submitted to any other Institution for any other degree/diploma/certificate in this university or any other University of India or abroad.

III. We have followed the guidelines provided by the university in writing the report.

IV. Whenever we have used materials (data, theoretical analysis, and text) from other sources, we have given due credit to them in the text of the report and giving their details in the reference.

Parangat Narsingh Pradhan (2018009122)

Rijesh Shrestha (2018011897)

**II**

**CERTIFICATE OF GUIDE**

This is to certify that this project report entitled “Dog Breed Recognition” submitted to Sharda University, is a Bonafide record of work done by Parangat Narsingh Pradhan and Rijesh Shrestha under my supervision from the start of the project to the end.

Signature of the Supervisor

Kanika Singla

Assistant Professor

Sharda University

**III**

**ABSTRACT**

In this project, we attempted to build a classifier capable of identifying a dog’s breed through a photograph. Many dogs of different breeds of dogs are very difficult to identify and classify even for humans. For training and evaluating the model we have used the Stanford Dogs dataset which contains images of 120 breeds of dogs from around the world. This dataset has been built using images and annotation from ImageNet for the task of fine-grained image categorization. It consists of 20580 images of dogs. Our system consists of three major stages: facial KeyPoint localization, facial normalization, and breed classification. This paper describes how we have used deep learning through convolutional neural networks and transfer learning for identifying dog breeds. The main reason why we have done this project is that different dogs need to be rescued and medicated based according to their breeds and also identifying dog breeds that are not suitable for family.

Index terms: Dog breed Identification, Stanford Dataset, facial KeyPoint localization, facial normalization, breed classification, deep learning, convolutional neural networks, transfer learning

**IV**

**ACKNOWLEDGEMENT**

We would like to express our special thanks of gratitude to our Faculty Guide Ms. Kanika Singla who gave us the golden opportunity to showcase our skill by this wonderful project on the topic Dog Breed Recognition, which also helped us in doing a lot of Research and we came to know about so many new things, we are really thankful to her.

**V**

**LIST OF FIGURES**

1. Fig1: Alaskan Malamute vs. Siberian Husky
2. Fig2: Boston Terrier vs. French Bulldog
3. Fig 3: Basic work flow of our project
4. Fig 4&5: Qualitative results of our KeyPoint detection
5. Fig 6&7: Our model identifying breeds of dog through CNN
6. Fig 8: Real Neural Network VS Artificial Convolutional Neural Network
7. Fig 9: Our model identifying the wrong breed of dog
8. Fig 10: Overview of our Proposed System
9. Fig 11: Sample images of our dataset
10. Fig 12: Examples of augmented images
11. Fig 13: Detecting all the Facial Key points
12. Fig 14: Examples of images before and after facial normalization. Concentrating on the face eliminates a huge amount of noise in the image and allows models to focus on a static, rigid object.
13. Fig 15: Identification of Dog’s breed
14. Fig 16: flowchart showing the step-by-step process of the project
15. Fig 17: DFD showing image processing

**VI**

**LIST OF TABLES**

1. Table 1: Different ways for obtaining System Development

**VII**

**1. INTRODUCTION**

**1.1 PROBLEM DEFINITION:**

The great variety of dogs poses a significant problem to those who would be interested in acquiring a new canine companion. However, walking down the street or sitting in somewhere, one might see a friendly, attractive dog and wonder of its breed. In many situations, it is impossible to ask an owner about the breed and in many cases, the owner themselves will be either unsure or incorrect in their assessment. It is also important for identifying the breeds of dog while rescuing them and also medicating them. Some breeds may also be highly aggressive and not suitable for family companion, but useful for police and army. This project is useful for identifying people’s ideal companion and for rescuing dogs.

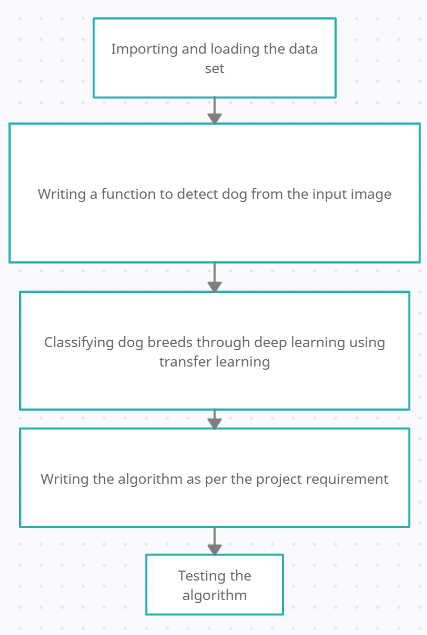
 **Fig1: Alaskan Malamute vs. Siberian Husky**

 **Fig2: Boston Terrier vs. French Bulldog**

**1**

**1.2 PROJECT OVERVIEW:**

The main goal is to implement a deep learning network to classify dog breeds from images. The goal is to evaluate different CNN custom architectures to solve this problem. First the image of the dog is taken as an input data. Then the input data is converted into blacks and white images. The data is broken into three sets which are basically training set, testing set, and validation set. The training dataset is used for calculating the pattern in images and adjusting the images. This training dataset is passed through a deep neural network. Feedforward is done and backpropagation is done to calculate the error and adjustment of weights. Then we plug in the image to know the breed of a dog. This breed is then revealed by feedforward. The probabilities of dog’s breed are predicted and thus output is calculated.



**Fig 3: Basic work flow of our project**

Fig3 shows the basic architecture/work flow of our project. The different steps which we have taken in our project are given below

1. The dataset is imported from the web and is broken into three parts i.e., training, validation and testing.

2. The dog images are converted into black and white because we don’t need any color in our case to find out the pattern.

3. We have 120 breeds so we have to use one hot encoding and set the position of the data. The breed corresponding to the input is set to one and others are set to zero.

4. We have imported the bottleneck features of xception model for dog breed classification.

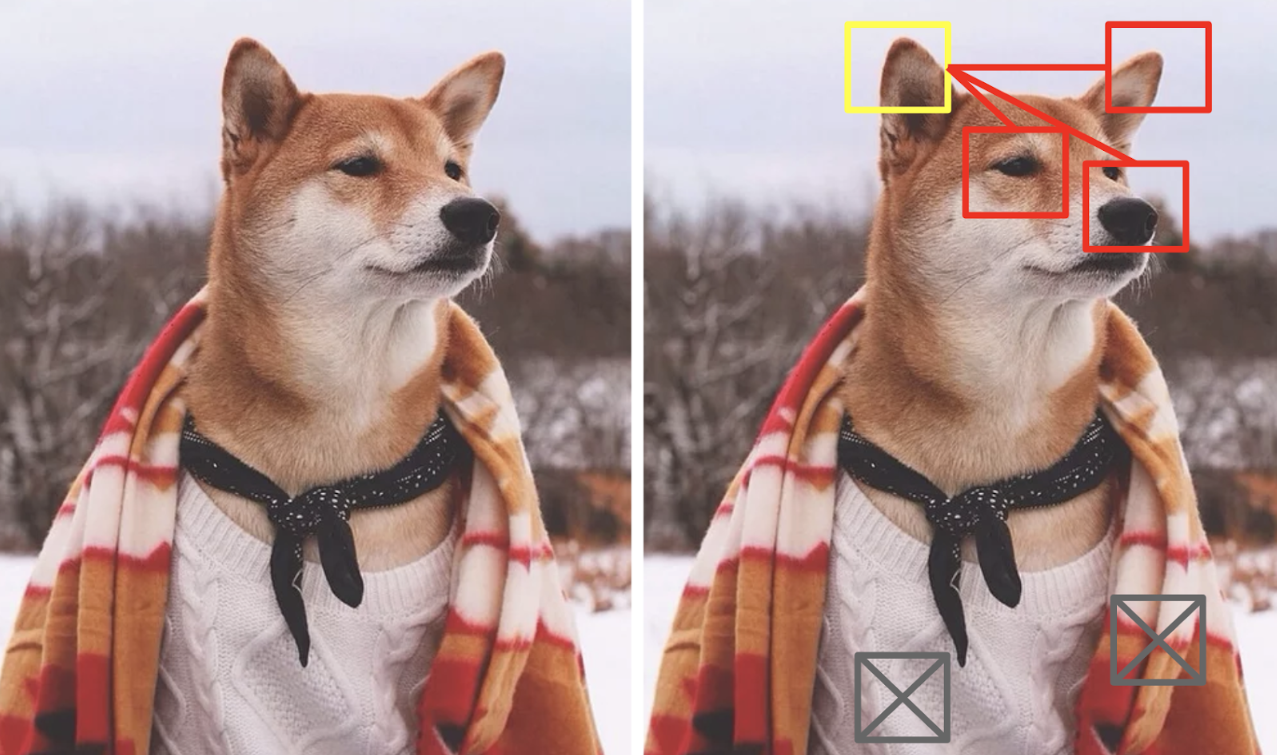
5. We have built our model and added in front of the pre trained xception model.

6. Now based on the training dataset and validation dataset the model is trained and validated. The weights are adjusted based on back propagation.

7. Now the dataset is tested on testing dataset and accuracy is computed.

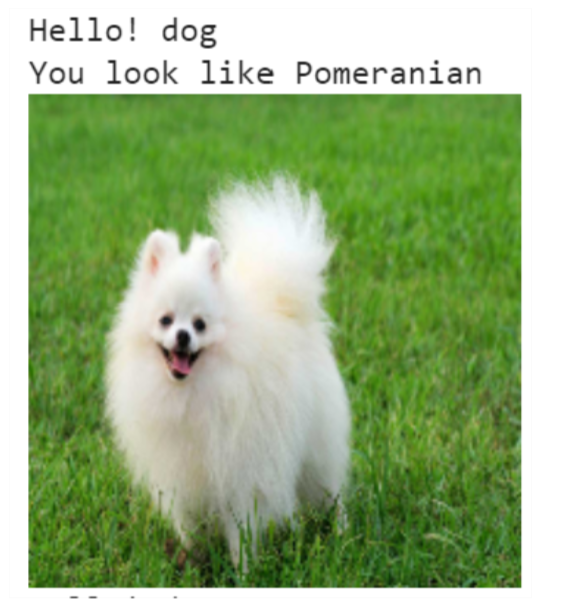
8. We plug in any image of the dog and check for the accuracy.

Models such as Convolutional Neural Networks (CNN) allow computers to automatically extract hierarchies of features from raw pixels. These techniques proved to be successful in a variety of visual analysis tasks. In this work CNN - based approaches were applied to the problem of dog breed identification. A number of techniques for improving performance of the model were used. Next, we experimented with data augmentation and transfer learning - two techniques that can potentially achieve high performance on small datasets. In this work it was shown that the models that use these two techniques can perform very well on a dog identification task.

**Fig4****Fig5**

**Fig 4&5: Qualitative results of our KeyPoint detection**

**Fig6**

**Fig7**

**Fig 6&7: Our model identifying breeds of dog through CNN**

**2**

**1.3 HARDWARE REQUIREMENTS:**

* All the physical equipment’s i.e., input devices, processor, and output device & inter connecting processor of the computer s called as hardware.
* Hard Disk minimum of 40 GB.
* RAM minimum of 2 GB.
* Dual Core and up ,15” Monitor.
* Integrated webcam or external webcam (15 -20 fps).

**1.4 SOFTWARE REQUIREMENTS:**

* Operating system- Microsoft Windows 7 SP 1 or above
* Microsoft Visual Studio 2010
* MinGW and Visual C++ compilers (for Windows)
* Supporting Webcam Drivers

**3**

**2. LITERATURE SURVEY**

**2.1 BASIC DEFINITONS:**

**What is Machine Learning?**

Machine learning is a field of Artificial Intelligence that uses data analysis and machine learning algorithms to create a model and trains itself according to the training data and learns some parameters according to it and uses them to predict some output according to those parameters. The data provided is divided to training and testing data and the accuracy of the algorithm is calculated on the basis of test score it obtains. Machine Learning is mainly divided into 3 categories:

* **Supervised Machine Learning:**

In this type of algorithm, the training data is used to analyze the data given, learns some parameters and creates a function which takes the data entries as input and passes them through that function to calculate the output. In input data the features and output columns are clearly separated. Almost 70% of the machine learning is Supervised Machine Learning. Algorithms which use supervised machine learning are:

Linear Regression, Logistic Regression, KNN, K-Means, Random Forest, Naïve Bays

* **Unsupervised Machine Learning:**

This kind of algorithm is used when the input data is not classified as well as not labeled. These algorithms identify the patterns in the provided data and draw inferences from the data sets to identify the data which is unstructured and unlabeled. However, these algorithms are not capable of predicting very accurate results but still it provides quite efficient results base on the kind of unstructured data provided. These algorithms are more complex and difficult to understand than the supervised algorithms. Algorithms which use unsupervised learning are:

Hierarchical clustering, K-means clustering, Expectation minimization

* **Semi-supervised Machine Learning:**

These are the algorithms which use the combination of both supervised and unsupervised learning techniques. They use both labeled as well as unlabeled data to compute the patters available and create function to predict the output for the rest data. For its working it can use a bit of labeled data and most part of unlabeled data or its vice versa. Classification, Regression and recognition are some of the methods it uses. Algorithms which use unsupervised learning are:

Face recognition techniques, Voice recognition techniques

* **Reinforcement Learning Algorithm:**

It is a kind of learning where the model interacts with the real-world environment for training itself and discover patterns. Here the model learns from its own outputs. Trial and error method is applied where a model performs a lot of trials on the data and tries to predict the output based on previous learning, and for every correct prediction the model is given some positive reward and for a wrong prediction it is given some negative points. Hence by making more and more trials the accuracy of the model increases significantly. Reward and feedback are required by the agent to learn new features and increase accuracy of the system. Algorithms which use Reinforcement learning are:

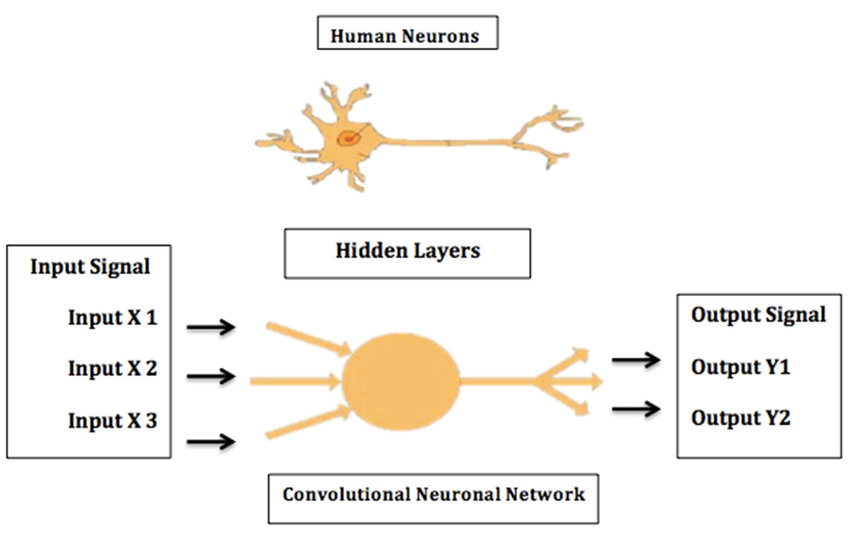
Q Learning Algorithm, State-Action-Reward-State-Action (SARSA), Deep Q Network (DQN), Deep Deterministic Policy Gradient (DDPG)

**What is Deep Learning?**

Deep learning is an Artificial Intelligence (AI) function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of Machine Learning in artificial intelligence that has networks capable of learning unsupervised from data that is unstructured or unlabeled. It is also known as deep neural learning or deep neural network.

**What is Convolutional Neural Network (CNN)?**

A Convolutional Neural Network (CNN) is a type of artificial Neural Network used in image recognition and processing that is specially designed to process pixel data. It is a spatial neural network which is mostly used to recognize patterns in an image as an input. CNN imitates the human brain. Our brain is a vast collection of Neurons. Convolutional Neural Network consists of neurons, weights and biases. This network receives the input and uses the biases and weights to calculate the weighted sum of the input and passes this sum through an activation function and finally comes up with a prediction or an output.



**Fig 8:** **Real Neural Network VS Artificial Convolutional Neural Network**

CNN consists of the following three layers:

* **Convolution layer of CNN:**

Convolutional layers are the layers where filters are applied to the original image, or to other feature maps in a deep CNN. This is where most of the user-specified parameters are in the network. The most important parameters are the number of kernels and the size of the kernels. Convolutional layers apply a convolution operation to the input, passing the result to the next layer. A convolution converts all the pixels in its receptive field into a single value. The final output of the convolutional layer is a vector.

* **Pooling layer** **of CNN:**

Pooling layers reduce the dimensions of data by combining the outputs of neuron clusters at one layer into a single neuron in the next layer. Local pooling combines small clusters, tiling sizes such as 2 x 2 are commonly used. Global pooling acts on all the neurons of the feature map. It is mainly used when the size of the image is too large. If the size of the image is too large, we need to reduce the number of learning parameters present in that image recognition. And while doing so it is required to insert pooling layers in between them.

The 2 famous types of Pooling Layers are as follows:

1. **Max Pooling:**

Max pooling is a pooling operation that selects the maximum element from the region of the feature map covered by the filter.

**2. Average Pooling:**

Average pooling computes the average of the elements present in the region of feature map covered by the filter.

* **Output layer of CNN:**

The output layer in a CNN is a fully connected layer, where the input from the other layers is flattened and sent so as the transform the output into the number of classes as desired by the network. The output is then generated through the output layer and is compared to the output layer for error generation. A loss function is defined in the fully connected output layer to compute the mean square loss. The gradient of error is then calculated. The error is then backpropagated to update the weights and bias values. One training cycle is completed in a single forward and backward pass.

**1**

**2.1 EXISTING SYSTEM:**

As the Convolutional Neural Network (CNN) is currently being trained the output accuracy is not very high and also the model is still slow. We are still adding the images of dogs manually but once the model or Convolutional Neural Network (CNN) is completely trained we are aiming for an output accuracy of about 90% to 95%. Currently our model predicts the breed of the dog from the image in the first 10 tries with an accuracy of 55%.

Some of the disadvantages of our existing system are as follows:

* Very time consuming.
* Highly error prone
* Lot of paper work results in lot of confusion.
* Updating and Retrieval tasks are very tedious.

To avoid all these limitations and make the system working more accurately it needs to be Computerized.



**Fig 9: Our model identifying the wrong breed of dog**

The table below shows the ways of improving dog breed identification based on previous research

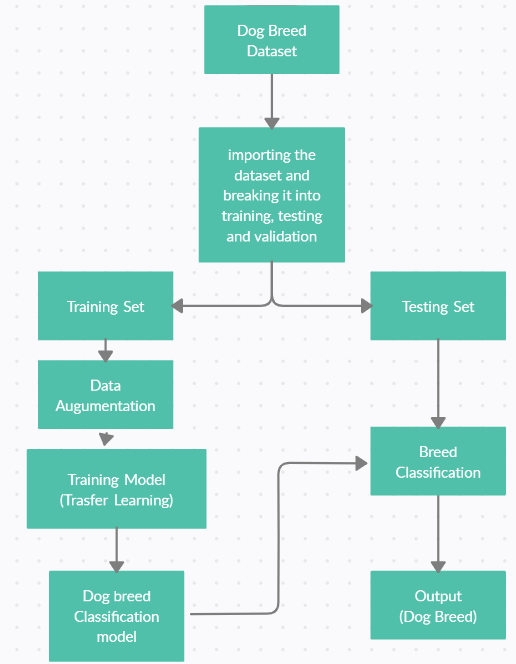
|  |  |
| --- | --- |
| **LITERATURE TITLE** | **LITERATURE REVIWW** |
| Dog Breed Classification Using Part Localization  (Liu, 2012)  Going Deeper with Convolutions  (Christian, 2014) | The performance of fine-grained classification can be improved by using part localization since dog breeds are similar in common parts but different in shape and appearance. |
| Dog Breed Identification  (LaRow, 2016) | The team picked the best model after comparing the accuracy of different machine learning models. |
| Transfer Learning for Image Classification of Various Dog Breeds  (Devikar, 2016) | Image Classification in CNN has proven to be highly efficient, but it requires a large training data set and substantial time for training to achieve higher accuracy. |
| TensorFlow, A System for Large-scale Machine Learning  (Abadi, 2016) | Using larger internal cluster with GPU can lead to fewer steps and high accuracy of the Inception model. |
| A Case Study on TensorFlow and Artificial Neural Networks  (Vivekanandan, 2017) | TensorFlow performs very well in recognition problems and the performance can be further improved by having more iterations. |

**Table 1: Different ways for obtaining System Development**

**2**

**2.2 PROPOSED SYSTEM:**

In this section, our proposed method for Dog Breed Identification is discussed in details. The proposed system is summarized in Fig. 10. It consists of 3 main steps importing the data set, training the dataset and testing the data set. The output from the training model is a dog breed identification model.



**Fig 10: Overview of our Proposed System**

Some of the advantages of our Proposed System are as follows:

* Greater efficiency.
* User friendly and interactive.
* Saves lot of time.
* Ensure data accuracy

**3**

**2.3 FEASIBILITY STUDY:**

Image recognition and classification have successfully applied in various domains, such as face recognition and scene understanding for autonomous driving. At present, human face identification is successfully used for authentication and security purposes in many applications. Therefore, there are attempts to extend studies from human to animal recognition. In particular, dogs are one of the most common animals. Since there are more than 180 dog breeds, dog breed recognition can be an essential task in order to provide proper training and health treatment. Previously, dog breed recognition is done by human experts. However, some dog breeds might be challenging to evaluate due to the lack of experts and the difficulty of breeds' patterns themselves. It also takes time for each evaluation. In this study, we use a public dataset to evaluate our method. The Stanford Dogs Dataset and Columbia Dogs Dataset are the public datasets for dog breed classification. We employ the Stanford Dogs Dataset as the data in this study. The Stanford Dogs dataset contains images of 120 breeds of dogs from around the world. This dataset has been built using images and annotation from ImageNet for the task of fine-grained image categorization. Contents of this dataset. It consists of around 20580 images of dogs. The sample images are shown in fig 11. Given the original images, it requires some pre-processing such as cropping and rescaling to extract dog faces as shown in Fig. 12. The pre-processed data is then split into a training set and testing set. The training set is augmented using data wrapping techniques such as rotation, flipping and adding noise. In this project, the dog breed classification model is constructed by using transfer learning techniques.

**Fig 11: Sample images of our dataset**



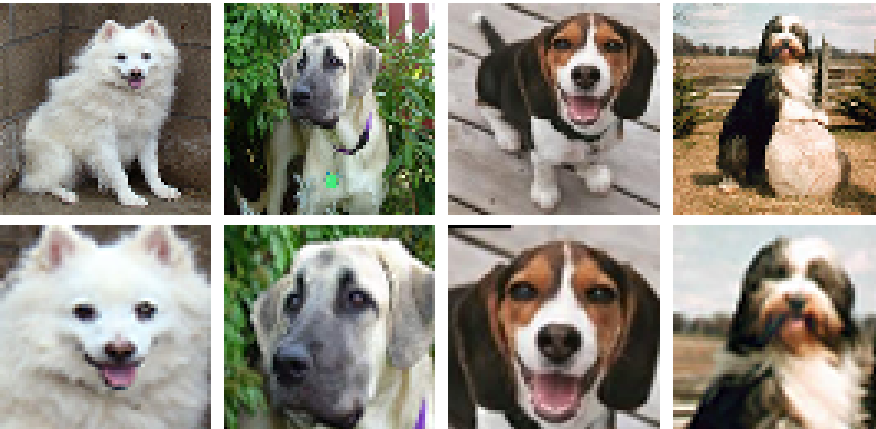
**Fig 12: Examples of augmented images**

Our dog breed identification algorithm also heavily relies on accurate facial detection as an initial step. We trained a Convolutional Neural Network (CNN) with a mean squared error loss function on all the facial Key Points. The output of the prototype was satisfying so we went through with this model.



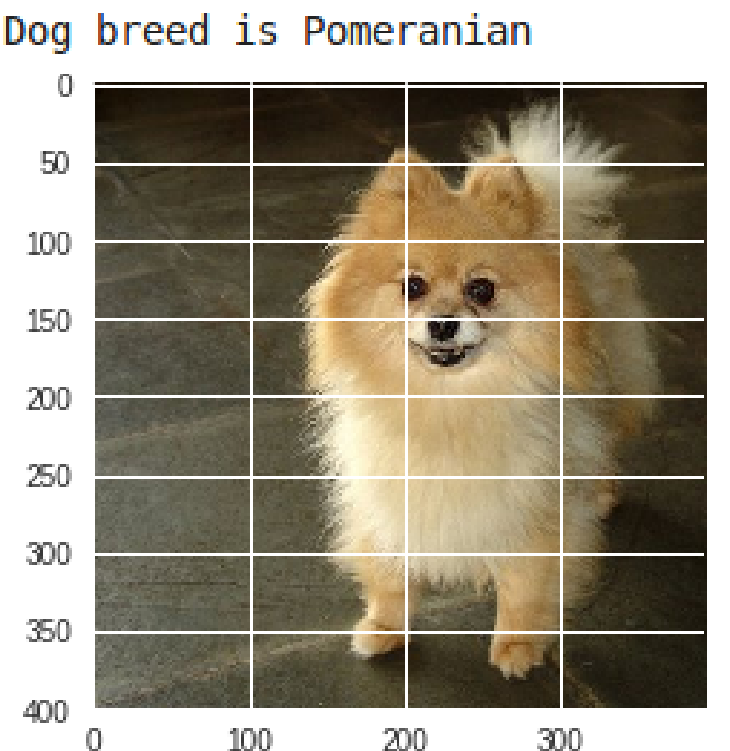
**Fig 13: Detecting all the Facial Key points**

Once the facial KeyPoint locations have been predicted by the first phase of the breed identification pipeline, normalized sub images of a constant size are prepared for the second neural network. This step is extremely important for the performance of any subsequent analysis. First, the center of the face is estimated as the mode of the midpoint between the eyes and the nose. Next, the slope of the vector from the left eye to the right eye is calculated and the entire image is rotated so that it lies flat. Finally, the length of the segment between the eyes, is calculated and a box centered at the center of the face with side length four times the interocular distance is cropped from the image. This box is then scaled to a constant size to serve as input to the next phase of the pipeline, breed identification. Before and after examples of this process are included in Fig 14.



**Fig 14: Examples of images before and after facial normalization. Concentrating on the face eliminates a huge amount of noise in the image and allows models to focus on a static, rigid object.**

Once facial Normalization is used on the input image. The image is run through the Dataset which we have uploaded and the CNN identifies the breed of the dog in the first ten tries with an accuracy of 55%. When the CNN is completely trained, we are aiming for an accuracy of about 90% to 95%.



**Fig 15: Identification of Dog’s breed**

**4**

**3. SYSTEM ANALYESIS & DESIGN**

**3.1 REQUIREMENT SPECIFICATION:**

Functional requirements specify the main technical functionalities and specifications that the system should incorporate.

**3.1.1 FACIAL KEYPOINT LOCALIZATION:**

Once the input image is received this software is used to detect all the facial key points in the dog’s face. The facial key points are located using Convolutional Neural Network (CNN). The facial key points in a dog which are needed for identification are the eyes, nose, bottom part of ear, top part of ear, and the top of the head. In total there are 8 facial key points.

**3.1.2 FACIAL NORMALIZATION:**

Once Localization is done then the unnecessary parts of the image are removed by finding the distance of the facial key points in simple words. Then the image is rotated and translated to filter out the noise. Once everything is done the image is run through the database where the CNN identifies the breed of the dog.

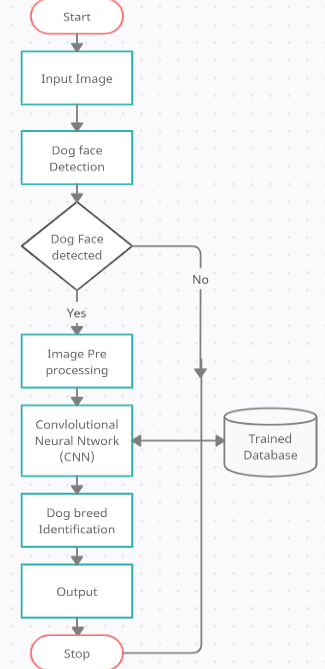
**3.1.3 TRANSFER LEARNING:**

In deep learning, transfer learning is a technique whereby a neural network model is first trained on a problem similar to the problem that is being solved. Transfer learning has the advantage of decreasing the training time for a learning model and can result in lower generalization error.

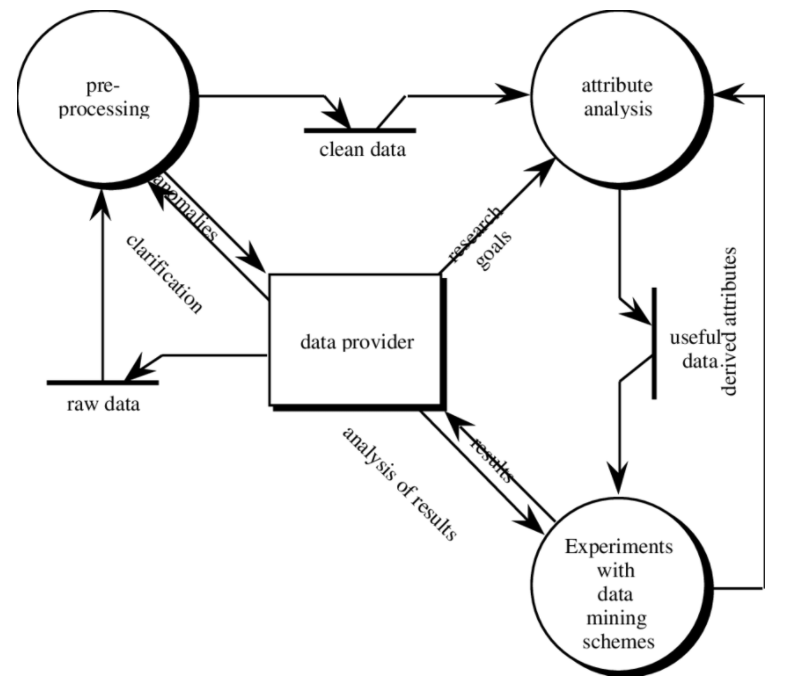
**1**

**3.2 FLOWCHART & DFD:**

Our project flow of how we have built and trained our model is shown in the flowchart and DFD below



**Fig 16: flowchart showing the step-by-step process of the project**



**Fig 17: DFD showing image processing**

**2**

**3.3 STEPS TAKEN:**

**3**

**3.4 ALGORITHMS USED:**

**4**

**4. RESULTS & OUTPUT 1**

**5. CONCLUSIONS**

In the end, we concluded that deep learning models have a very great capability to surpass the human potential if the data provided is sufficient. Engineers and scientists are still working on the deep learning field because till now the exploration of deep learning is limited. In the future, the deep learning will create another deep learning models on its own and deep learning model will write codes and surpass the human capabilities. Deep learning has a lot of scope in medical sciences by analyzing the images by deep convolution neural network. Deep learning is the future of technology and it will see many advancements in the field. Once deep learning has been developed it will also help in flourishing other sectors as well. Dog breed identifier is one of the mini projects of deep learning developed using advance neural networks. Transfer learning has a great scope in the future by combining a prebuilt model with the model we constructed. This Dog breed identification can also be used in Cat breed, Horse breed and Car model identification by adjusting the required data set and training the Convolutional Neural Network (CNN).

**2**

**6. REFERENCES**

1. Dog Breed Classifier Using Deep Learning

Agrim Dogra and Harsh Massand

2. Dog Breed Classification Using Part Localization

J. Liu, A. Kanazawa, D. Jacobs, and P. Belhumeur

3. Dog Identification using Soft Biometrics and Neural Networks

Kenneth Lai, Xinyuan Tu and Svetlana Yanushkevich

4. Dog breed classification via landmarks

X. Wang, V. Ly, S. Sorensen, and C. Kambhamettu

5. Automatic Dog Breed Identification

Dylan Rhodes

6. Using Convolutional Neural Networks to Classify Dog Breeds

Hsu, David

7. Knowing Your Dog Breed: Identifying a Dog Breed with Deep Learning

Punyanuch Borwarnginn, Worapan Kusakunniran, Sarattha Karnjanapreechakorn and Kittikhun Thongkanchorn

8. Dog Breed Identification

Whitney LaRow, Brian Mittl and Vijay Singh

9. Dog Breed Identification Using Deep Learning

Zalan Raduly, Csaba Sulyok, Zsolt Vadaszi and Attila Zolde

10. Dog Breed Identification

Yerbol Aussat

11. image classification for dogs and cats

K. Zhou, B. Liu and Y. Liu.

12. Dog Breed Identification with Fine tuning of pre-trained models

B. Vijaya Kumar and K. Bhavya

13. Transfer learning for image classification of various dog breeds

Pratik Devikar

14. Transfer learning on convolutional neural networks for dog identification

X. Y. Tu, K. Lai and S. Yanushkevich

15. Dog breed classification using part localization

J. L. et al

**3**