Automatic Classification of Breeds of Dog using Convolutional Neural Network

P.O. Adejumobi[1]*, I.O. Adejumobi[1], O.A. Adebisi[2], S.O. Ayanlade[3], I.I. Adeaga[4]

[1]Department of Electronic and Electrical Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.
[2]Department of Mechanical and Mechatronics Engineering, First Technical University, Ibadan, Nigeria.
[3]Department of Electrical and Electronic Engineering, Lead City University, Ibadan, Nigeria.
[4]Department of Computer Engineering, The Polytechnic Ibadan, Ibadan, Nigeria.

ABSTRACT: Dog is a mammal that has been a friend of man for ages, it is naturally a domestic animal with a high level of phenotype differences in behaviour and morphology. Breeding and crossbreeding activities have increased the number of dog breeds globally, thereby resulting in dogs with inter breed similarities and intra breed differences thereby creating a difficulty in their classification. The American Kennel Club (AKC) classified breeds of dog into groups based on characteristic, purpose, behaviuor and uses in order to optimize the potentials in the breeds. However, most people find it difficult to identify and classify the dog breed groups. Existing works did not consider the automatic grouping of dog breeds. Hence, there is need for automatic techniques to classify dog breeds into groups with improved accuracy. This work used the concept of Convolutional Neural Network (CNN) to develop a model that will automatically classify dog breeds into group based on the American Kennel Club standard using the Stanford's dog dataset. The developed model achieved 92.2% accuracy, 80.0% sensitivity, 95.3% specificity and 93.4% area under curve (AUC). The model's performance is excellent compared to existing works that used the same dataset. The experimental result was validated with two classic CNN models (ResNet-50 and SqueezeNet) using the same parameters.

KEYWORDS: convolutional neural network; deep learning; dog breeds; machine learning; Stanford's dog dataset

[Received Dec. 7, 2022; Revised June 30, 2023; Accepted July 29, 2023] Print ISSN: 0189-9546 | Online ISSN: 2437-2110

I. INTRODUCTION

Dog (Canis Familiaris) is a mammal that has been a friend of man for ages (Akash *et al*, 2021). It is a famous pet because it is generally friendly, playful, loyal and obedient (Mohamed *et al*, 2020). Dogs are products of human choice in behaviour which distinguished them from other domestic animals (Ayan *et al*, 2022). In addition, it is the most morphologically and genetically different species on earth (Whitney *et al*, 2015) with high level of phenotype differences in behaviour; resulting in its excellent performances at work and social roles such as security guards, hunting aides, fighters, artists, guides, garbage disposers, pets, source of food and fiber (James and Deborah, 2014). In general, dog has rapid adaptability to variable ecological locations to meet various social and instrumental demands of human beings (Kenth and Bjorn, 2012). Physical attributes such as size, acuity, quality, colour and coat length determine human choice and preferences for dogs.

The introduction of dog breeding in the 19th century enhanced the retention of phenotypic traits amongst breeds. Dogs are bred for definite hereditary trait that ranged from physical features to personality trait (Akash *et al*., 2021). These enabled researchers to acquire understanding regarding genetic bases, behavioural differences and domestication effects on the temperament of dog breeds (James and Deborah, 2014). Different dog breeds have specific characteristics that describes them. Breed diversity made dog the most interesting animals for experiment (Suyash *et al*, 2021), while some breeds differ in characteristics such as behaviour towards humans, likes, dislikes, popularity, size, shape (Akash *et al*, 2021) and peculiar health conditions (Suyash *et al*. 2021); others are fairly alike in body structures and features. While some breeds of dog have similarity in facial characteristics, they vary expressively in colour and those that share similarity in colour significantly vary in facial characteristics (Aditya *et al*, 2013). In addition, there exists a great intra-class variation amongst some dog breeds as they neither share a consistent color of fur nor sizes due to age. Thus, making their identification, categorization and differentiation a challenge.

The world's major body for dog breeds known as Federation Cynologique Internationale (FCI) has identified 340 breeds of dog while the American Kennel Club identified 192 dog breeds (Mohamed *et al*, 2020). At the moment, there exist over 450 breeds of dog. The continuous breeding and crossbreeding of dogs is progressively increasing the available number of breeds to meet the global demand for diverse breeds of dog for different applications (Ayan *et al*, 2022). Factors such as popularity, rarity and pedigree influence the prices of dog. Therefore, choosing a dog breed would involve considering factors such as cost, activity level and

*Corresponding author: adejumobiphilip@gmail.com
doi: http://dx.doi.org/10.4314/njtd.v18i4.1485

temperament (Ayan *et al*, 2022). Figure 1 shows the images of ten dog breeds.

Dog breed grouping is the classification of breeds based on similarities in appearance, specified purpose, temperament and ancestry (Kenth *et al*, 2012). A dog group contains a variety of dog breeds that vary dramatically in physical appearances but are similar in behaviour (Kenth *et al*, 2012). The American Kennel Club (AKC) conducted a comprehensive historical background analysis to identify dog group using the breed standard (Kenneth *et al*, 2019). In the experiment, the behaviour of 30 popular and registered dog breeds was used to categorize them into seven different groups. The groups are hound, terrier, toy, sporting, working, non-sporting and herding (James and Deborah, 2014). An extension of the dog breed group contains therapy dogs, support dogs and hybrid dogs (Kenneth *et al*, 2019). Dog breed group provides appropriate alternatives when making choices in terms of budget, specific tasks and areas of application.

kernel is sliding the input, the stride parameter is used to determine the number of pixels to skip. Stride value ranges from 1 to 3 depending upon the amount of loss which can be accommodated during convolution (Ravi and Shailender 2019).

This work aims to automatically classify dog breeds into groups based on the American Kennel Club's standard using a designed convolutional neural network with less computational requirements. The remaining part of this paper is as structured. Section II presents the related work; Section III presents the methodology. Section IV presents the results and discussion and section V presents the conclusion and future work.

II. LITERATURE REVIEW

The characteristics of each dog group is itemized as follows:

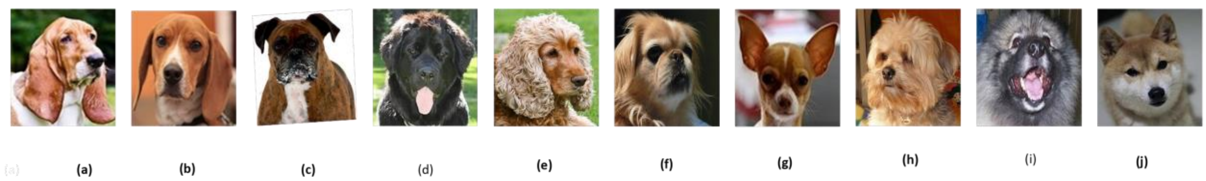

Figure 1: Illustration of dog breeds (a) Basset hound (b) Beagle (c) Boxer (d) Newfoundland (e) English cocker spaniel (f) Japanese chin (g) Havanese (h) Chihuahua (i) Keeshond (j) Shiba-inu

The identification of dog breeds and groups is a big challenge for most people (Amit *et al*, 2021). Dog breeds can be identified using an expert-based technique; in this method, dog experts identify different variety of dog breeds (Whitney *et al*, 2015). However, this method is limited due to the availability of few dog experts and the accuracy of the method is prone to human error (Akash *et al*, 2021). The introduction of DNA test provided an accurate and precise result however, the process is expensive and complex (Mohamed *et al*, 2020). Hence, there is need to develop less expensive and automatic methods for dog breed group identification and classification. The development of artificial intelligence has outperformed human capacity in image identification and classification tasks (Alexandre and Mauricio, 2017). Also, the introduction of deep learning techniques has provided a better alternative for image classification with excellent results (Ghirlanda *et al*, 2013). The automatic classification of dog breeds into group using Convolutional neural network (CNN) would be an interesting research area.

Convolutional Neural Network (CNN) is a feed-forward neural networks that can identify, classify and recognize features in an image (Adebisi *et al*., 2020). It has gained a lot of research attention due to its high capability to correctly learn features and classify images automatically. Each layer of the CNN has specific function it performs (Philip *et al*., 2023). The prime operation of convolutional layer is convolution operation (Alex *et al*., 2012) where the input is convolved with the filters to generate the output feature maps. Padding is used to preserve the size of the output, it increases the size of the input data through filling constants around input data. As the

Herding group: Dogs in this group are athletic, intelligent, trainable, diligent and perform best with lots of exercise. They were bred to herd cattle and sheep. Examples of dogs in this group are; Sheepdogs, Collies, Cattle dogs, Shepherds, Corgis, Belgian Tervuren and Belgian Malinois (Naufal *et al*, 2022). Hound group: Dogs in this group are intelligent, independent, sprinters and affectionate. They were initially bred for hunting. Examples in this group are; Saluki, Sight hounds, Afghan, Ibizan, Greyhound, Pharoah, Italian Greyhound, Whippet, Rhodesian Ridgeback, Fox Hounds, Beagle, Bloodhound, Basset Hound and Otterhound (Maria, 2015). Non-Sporting Group: Dogs in this group are smart, playful, curious, loyal, friendly and eager. They are companion animals with variation in terms of shape, size, coat types and colors. Examples in this group are; Boston Terrier, American Eskimo, Shar-Pei, Bulldogs, Dalmation, Chow Chow, Keeshond, Shiba-Inu, Poodle, Lhasa Apso, Tibetan Terrier (Enya *et al*, 2022).

Sporting group: Dogs in this group are trainable, loyal, happy, bright, eager, friendly, lovable and confident. They were formerly bred to help hunters in pointing, finding, retrieving and flushing game. They are also used as assistance dogs, therapy dogs and rescue/ search dogs. Examples in this group are Spaniels, Clumber, American Water, English Cocker, English Toy, Irish Water, English Springer, Sussex Tibetan, Japanese Chin and Welsh Spaniels (Naufal *et al*, 2022). Terrier Group: Dogs in this group are energetic, scrappy, predatory, independent, smart and playful. They were bred to hunt and kill vermin such as foxes, rats and weasels. Examples in this group are Bedlington, Airedale, Boston, Scottish, Bull, Russell and Staffordshire (Naufal *et al*, 2022).

Toy Group: Dogs in this group are fearless, sensitive, lively, alert, intelligent, snappy and affectionate. They are small dogs with a lot of spirit and strong traits. Examples in this group are Chihuahua, Cavalier King Charles, Havanese, Maltese, Italian Greyhound, Pekinese, Pomeranian, Pug, Toy Poodle, Shih-Tzu and Yorkshire Terriers (Maria, 2015). Working Group: Dogs in this group are courageous, strong, affectionate, loyal, calm, protective, smart, athletic and confident. They have medium to large sizes, they were initially bred as hunters, guardian and draft dogs. Examples in this group are Siberian Husky, Akita, Samoyed, American Husky, Newfoundland, St. Bernard, Norwegian Elkhound, Rottweiler, Finnish Spitz, Boxer, Giant Schnauzer and German Shepherd (Enya *et al*, 2022).

In literatures, researchers have made efforts in the classification of dog breeds. (Jiongxin *et al*, 2012) classified dog breed using part localization. The dataset was downloaded from Image-net, Google and Flickr. The training set contained 4,776 images while the test set contained 3,575 images. Amazon's Mechanical Turk (MTurk) was used for dog faces localization. The close visual similarity among breeds resulted in some of the errors obtained. Recognition rate of 67% was achieved in the experimental results which indicated that classification performance could be increased by accurate part localization in comparison to state-of-the-art techniques. The work also proved that extracting corresponding image parts such as face, eyes and breed-specific part localization can improve the classification performance. (Xiaolong *et al*, 2015) classified dog breed using landmarks on dataset gathered from Flicker and Image-Net online. 8, 351 images from 133 dissimilar dog breeds that made up the dataset. The geometry of a breed structure was described by Grassmann manifold while the dog faces geometry was categorized to differentiate dissimilar breeds by modelling characteristics of 2-D landmarks which were extracted from the breeds of dog. Experimental results indicated that the work outperformed state-of-the-art approaches by approximately 20%.

(Wenting *et al*, 2015) identified dog breed using four CNN models; AlexNet, DenseNet161, VGG16, and ResNet18. The dataset was downloaded from Kaggle and it contained 10,000 images from 120 breeds of dogs. There were 9,000 images in the training set and 1,000 images in the test set. Optimization was done to increase the identification accuracy of the models. Comparing the performances of the four models, DenseNet achieved the highest accuracy of 85.14%. (Whitney *et al*, 2015) performed dog breed identification using connected convolutional neural network. The training set contained 4,776 images while the test set contained 3,575 images. In the preprocessing, the images were scaled to be 128x128 pixels. OpenCVs feature detector and descriptor extractor were used to extract SIFT descriptors while bag of words model was used for the classification.

(Zalan *et al*, 2018) identified dog breed using deep learning models (NASNet-A mobile and Inception-Resnet V2) on the Stanford's dog dataset. There were 12,000 images in the training dataset while the test dataset contained 8,580 unevenly distributed images. In the preprocessing, the images of dog were resized into 299 x 299 pixels for (Inception-Resnet V2 input) and 256 x 256 pixels for (NASNet-A mobile). The system contained two main components: a centralized web server and a mobile client. NASNet-A mobile network was fine-tuned at an initial learning rate of 0.029 while the Inception-Resnet V2 was fine-tuned at an initial learning rate of 0.1. In the result, NASNet-A mobile architecture achieved 85.06% accuracy on the training dataset and 80.72% accuracy on the test dataset while Inception-Resnet V2 network achieved 93.66% accuracy on the training dataset and 90.69% accuracy on the test dataset. The developed system identified and gave detailed information of the input.

(Punyanuch *et al*, 2019) classified dog breeds using CNN with transfer learning. The work employed two approaches. The conventional based approach used Histogram Oriented Gradient (HOG) and Local Binary Pattern (LBP) while the second approach used deep learning with transfer machine. In the result, the retrained CNN model achieved 96.75% accuracy while the HOG descriptor achieved 79.25% accuracy. (Durga *et al*, 2019) developed an android application to determine dog breed from a snapped picture. The application was developed using Stanford's standard dog dataset on CNN pre-trained models; Inception-v3, Inception-ResNet-v2, VGG16 and Xception for the feature extraction. The dataset contained 9,199, 2,000 and 9,381 images respectively for the training, validation and test datasets. In the result, accuracy of 89%, 94%, 81% and 93% were obtained for Inception-v3, Inception-ResNet-v2, VGG16 and Xception respectively on the testing data.

(Bickey *et al*, 2020) classified breed of dog using convolutional neural network for facial recognition. The dataset contained 13,233 human images and 8,351 dog image. The work aimed at finding the percentage of dog features in human and human features in dogs. Similar features were stored in one group using principal component analysis while facial features were saved in vector form. In the experiment, features of the input dog were compared with the vector to output the most efficient result. In the result, whenever a dog image is the input; the algorithm outputs the dog breed and feature resemblance in the breed. On the other hand, if human image is supplied, the algorithm will determine the facial features of dog present in human and vice-versa. (Kanika *et al*, 2020) identified dog breeds using CNN architecture. Stanford Dogs dataset with four CNN models; Resnet101, Resnet50, InceptionResnetV2 and InceptionV3 were used for the experimentation. For the preprocessing, the resolution of the images varied with 4000x3000 pixels as higher resolution and 400x300 pixels as smaller resolution. The dataset contained a total of 20,579 images (10,222 images as training set and 10,357 images as test set). In the experimental result, the models achieved validation accuracy of 71.63%, 63.78%, 40.72% and 34.84% for Resnet101, Resnet50, InceptionResnetV2 and InceptionV3 respectively.

(Mohamed, 2020) identified dog breeds using Inception-Resnet V2 model on the Stanford's dog dataset. The work contained two segments; image processing using neural network and data rendering with web scraping. The data from different websites were collated and rendered in the application using web scraping technique. (Sneha *et al*, 2020) performed dog breed prediction using convolutional neural network. The Columbia dog dataset containing 8,350 images from 133 different dog breeds was used. The left and right eye, nose, left and right ear base, left and right ear tip, and head top were the marked facial key points. Each of classifiers was ran using SIFT descriptor feature set and the accuracy of each neural network model was compared. (Akash *et al*, 2021) classified dog breed using deep learning on Stanford dog dataset which contained 2,050 images from 120 categories of dog breeds. The work compared the performance of Inception V3 and VGG-16 models in the classification of dog breeds using the same parameters. In the result, Inception V3 achieved an accuracy of 85% while VGG-16 achieved an accuracy of 69%.

(Sandra and Remya, 2021) classified different dog breeds using CNN. The method employed deep learning innovative strategy, convolutional neural networks and transfer learning on two different datasets. The work achieved an accuracy of 90.86% and 93.53% on both datasets respectively. (Punyanuch *et al*, 2021) proposed deep learning approach to identify dog breed. Published dataset containing 133 dog breeds was used. The method identified dog breed using their face images. Three existing CNN models were pre-trained using transfer learning technique. The model achieved an accuracy of 89.92%. (Suyash *et al*, 2021) classified dog breed using convolutional neural network. The paper proposed a system that identified and classified various dog breeds irrespective of their age. The images were resized to 224 x 224 pixels, preprocessed and filtered. A CNN architecture was created from scratch with 3 convolutional layers, relu activation function and max pooling along with dropout for the dimensionality and feature reduction. The dataset was separated into training, validation and test. Features such as face structure, eye, hair and mouth were extracted using the designed model while the test dataset was tested using VGG-16 model. (Ayan *et al*, 2022) classified dog breeds using modified-xception model. The work built a convolutional neural network to achieve higher accuracy. The overall classification performance of the method was evaluated using the Kaggle dog breed identification dataset. The model achieved 87.40% accuracy.

Considering the reviewed works, most of the literatures reviewed focused on the classification of dog breeds, none of the work considered the classification of dog breeds into groups to optimize their potentials. Hence, this work aimed to classify dog breeds into group based on the American Kennel Club standard on a benchmark dataset with improved accuracy.

III. METHODOLOGY

A. Building CNN model

Researchers have developed various convolutional neural network models with varying architecture and depth to improve classification performance. However, these architectures require huge computational resources, requirements and dataset to be able to generalize well. Hence, there is need to develop CNN architectures with less computational requirement and dataset that will attain the performance of the existing classical models. CNN architecture can be designed by stacking convolutional layer, pooling layer and fully connected layer on each other until an acceptable accuracy is achieved (Muhammad *et al*. 2019). In the design of CNN architecture (Rikiya *et al*., 2018), a simple principle is that the feature space must be shallow and wide at the start of the network while it must be deeper and narrower at the end of the network (Manoj *et al*., 2018). Also, at a start of the network; smaller number of filters and channels must be used and progressively increased as the network increases (Neda, 2021). The hyperparameters for defining the layers in convolutional networks are filter, padding and stride. The filter fits weights through training and learning (Wenzhong, 2020).

We defined our proposed model architecture using convolutional layer, batch normalization layer, pooling layer, and fully connected layer. We initialized this network layers to create a model using sequential constructor. Layer-1 consists of the convolutional layer, layer-2 is the batch normalization layer, layer-3 is the rectified linear unit layer while layer-4 is the max pooling layer. Layer-5 consists the second convolutional layer, layer-6 is the batch normalization, layer-7 is the rectified linear unit layer while layer 8 is the max pooling layer. Layers 9 and 10 are fully connected layers. The size of the input image for our proposed model is 224 x 224 x 3. The size of the output feature map from each layer was computed using Eqn. (1).

$$O = \frac{W-K+2P}{S} + 1 \quad (1)$$

where: O is the output height/length, W is the input height/length, K is the filter size, P is the padding and S is the stride. The padding is determined using Eqn. (2).

$$P = \frac{K-1}{2} \quad (2)$$

where P denotes the filter padding, K denotes the filter size. Weights and bias are parameters in every layer of the convolutional neural network, they are calculated using Eqns. (3), (4) and (5).

$$W = K^2 \times C \times N \quad (3)$$

$$B = N \quad (4)$$

$$P = W + B \quad (5)$$

where W equals the number of weights, B is the number of bias, N is the number of kernels, C is the number of channels in the input image, K is the size of the filter while P is the number of parameters in a layer.

Figure 2(a) shows the flow chat for building CNN model while Figure 2(b) shows the flow chat for training CNN model.

Figure 3 shows the architecture of the proposed CNN model.

From Figure 3, Conv1 and Conv2 represents the convolutional layers, Pool1 and Pool2 represents the pooling layers while FC3 represents the fully connected layer. The built architecture comprises of two fully connected layers and two convolutional layers. The convolutional layers are followed by a batch normalization, ReLU and max pooling layers correspondingly. The first fully connected layer contains 512 neurons while the second fully connected layer has 5 neurons to compute the class scores. Each layer has its own hyper parameters, weights and biases. To achieve model’s flexibility and simplicity for our model, the CNN architecture was implemented using the Pytorch library in Python programming. The proposed CNN model configuration is provided in Table 1.

There are essential stages in the classification of dog breeds using convolutional neural network. The stages are image acquisition, image preprocessing, feature extraction and classification. Figure 4 shows the work flow diagram for dog breed classification.

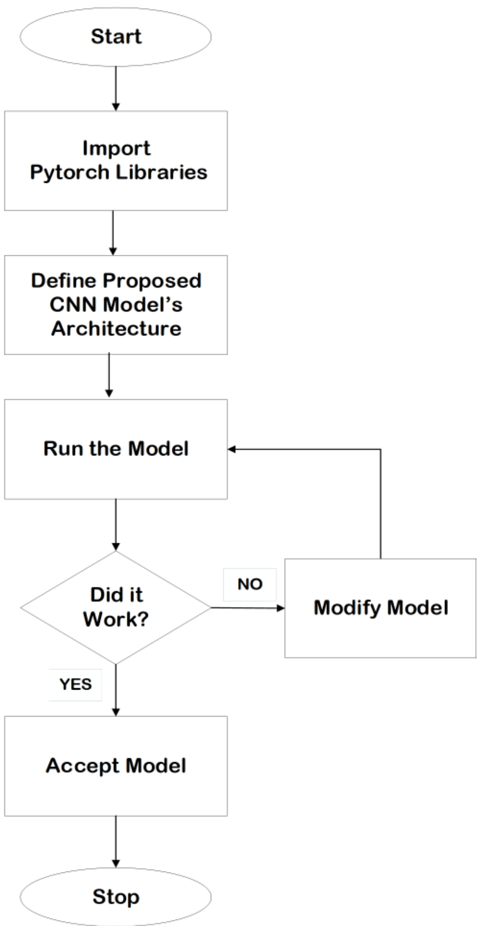

Figure 2(a): Flow chat for building CNN model

B. Dataset

The feat of deep learning models in image classification depends on the size of dataset (Althnian *et al*, 2021), the dataset is an important factor that determines the performance of convolutional neural network models. For machine learning algorithms, the data is divided into training dataset, validation dataset and test dataset. Data division configuration determines the performance of deep learning models. There is no fixed rule of partition ratio for the training and validation datasets, it depends on the size of the available dataset. However, the training dataset must be more than the test dataset. For very large dataset the ratio of training to validation dataset could be 80:20 or 90:10 while for small dataset the ratio could be 60:40, 65:35 or 70:30 (Althnian *et al*, 2021). The training dataset is used for training the model while the validation dataset is used to access the model's performance during the training. Hence, the model's performance is optimized by the training and validation dataset. The test dataset is used for testing the model's generalization performance, it is always concealed from the training and evaluation process so that the model could generalize well to unseen data (Aditya *et al*. 2013).

Stanford's dog dataset is part of the ImageNet challenging datasets (Aditya *et al*. 2013), this dataset is made up of 120 classes of dog breeds across the world. The Stanford dog

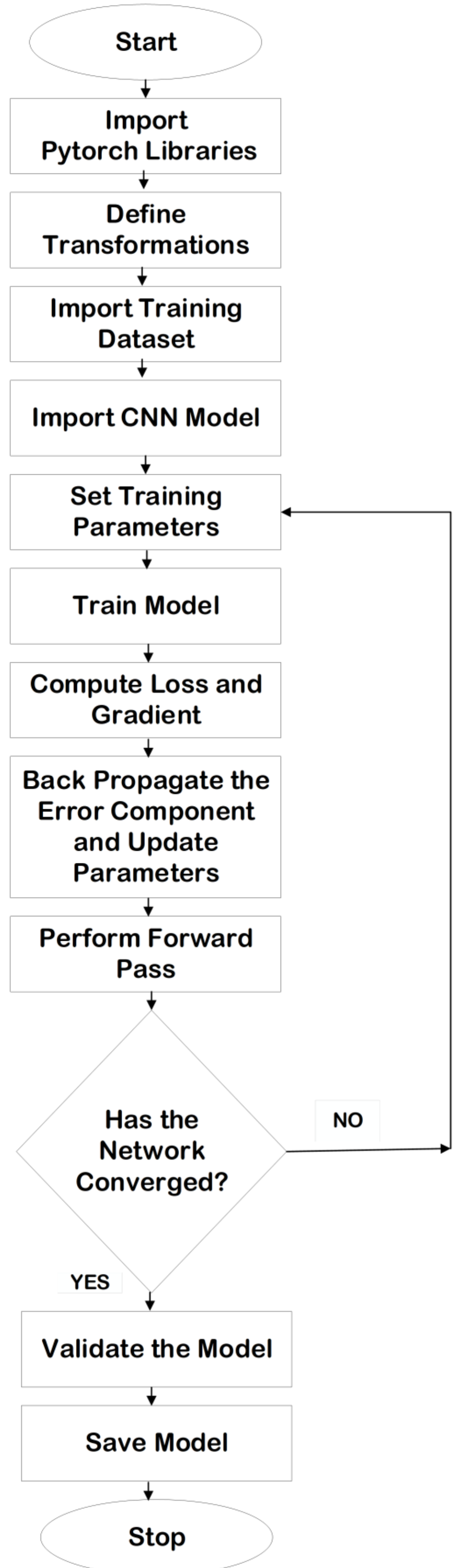

Figure 2(b) Flow chat for training CNN model

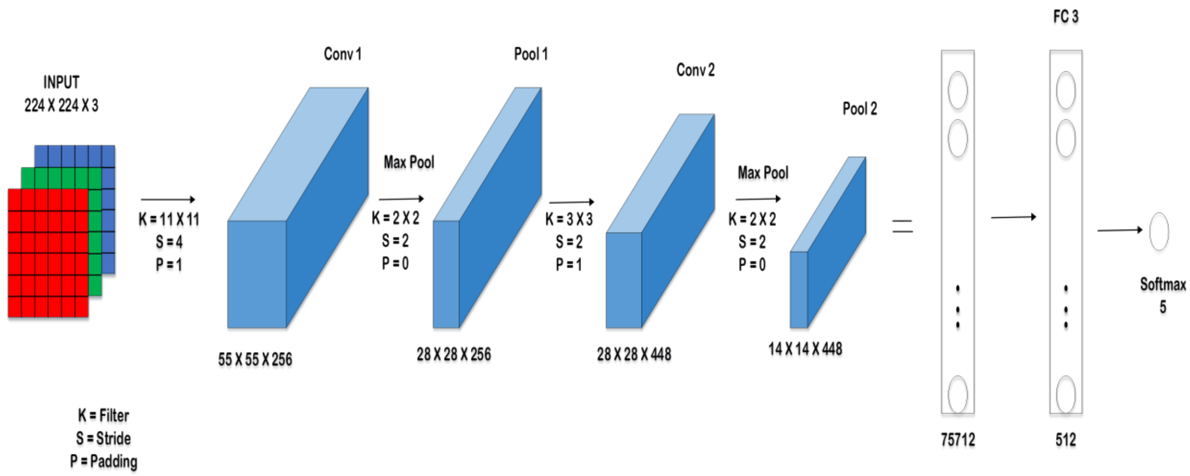

Figure 3: Architecture of the proposed CNN model

Table 1 Configuration of the proposed CNN model

Layer	Type	Size of input feature map	Dimension of filter	Depth of layer	Stride	padding	Size of output feature map	Parameters
0	Input	224x224x3	-	-	-	-	-	0
1	Convolutional	224x224x3	11x11	256	4	1	55x55x256	93184
2	Batch Normalization	-	-	256	-	-	55x55x256	1024
3	ReLU	-	-	-	-	-	-	0
4	Max-Pooling	55x55x256	2x2	-	2	-	28x28x256	0
5	Convolutional	28x28x256	3x3	448	1	1	28x28x448	101824
6	Batch Normalization	-	-	448	-	-	28x28x448	115136
7	ReLU	-	-	-	-	-	-	0
8	Max-Pooling	28x28x448	2x2	-	2	-	14x14x448	0
9	Fully Connected	-	-	512	-	-	512	44958208
10	Softmax	-	-	5	-	-	-	2565

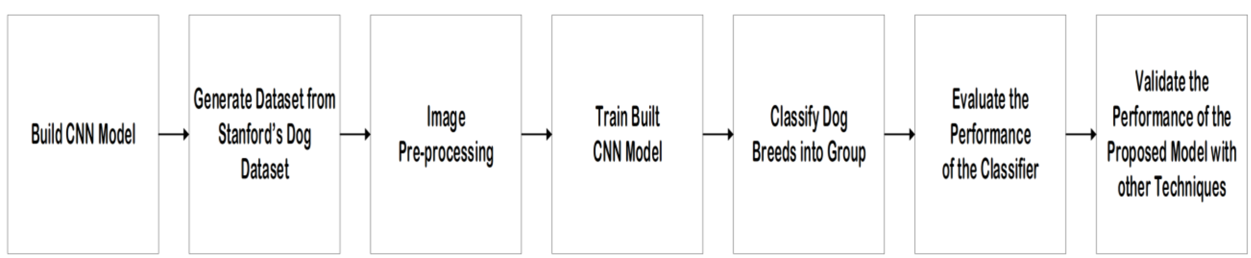

Figure 4: Flow diagram for dog breed classification

dataset contained between 150 to 200 images for each dog breed at different ages. There are 10,222 images in the training set and 10,357 images in the test set. In comparison with other animal datasets, the images in Stanford's dog dataset was captured in real-time as natural scenes leading to variations in image background. For this study, samples of ten recognized dog breeds (Basset hound, Beagle, Boxer, Newfoundland, English cocker spaniel, Japanese chin, Havanese, Chihuahua, Keeshond and Shiba-inu) were extracted from the Stanford's dog dataset to create our database.

C. *Image Preprocessing*

Image pre-processing is used to enhance the quality of images to achieve the best result. For this work, all duplicates in the extracted images for each class were removed. Then, the images were de-noised using median filtering. In other to overcome the difficulty of limited data, data augmentation was applied on the images using image transformation operations such as zooming, rotating the images randomly at 90° and 180° and flipping the images randomly vertically and horizontally. For this work, our database contained 550 images of Basset hound, 550 images of Beagle, 550 images of Boxer, 550 images of Newfoundland, 550 images of English cocker spaniel, 550 images of Japanese chin, 550 images of Havanese, 550 images of Chihuahua, 550 images of Keeshond, 550 images of Shiba-inu. Based on the American Kennel Club dog group classification, these ten breeds of dogs were grouped into five by integrating them into folders.

For a balanced dataset, 50 images were extracted from each breed class to generate 500 images (~10% data) for the test set; Hound-group (50 images of Basset hound and 50 images of Beagle), Non-sporting-group (50 images of Keeshond and 50 images of Shiba-inu), Sporting-group (50 images of English cocker spaniel and 50 images of Japanese chin), Toy-group (50 images of Havanese and 50 images of Chihuahua), Working-group (50 images of Boxer and 50 images of Newfoundland). The remaining 500 images per class were used for the training set as follows; Hound-group (500 images of Basset hound and 500 images of Beagle), Non-sporting-group (500 images of Keeshond and 500 images of Shiba-inu), Sporting-group (500 images of English cocker spaniel and 500 images of Japanese chin), Toy-group (500 images of Havanese and 500 images of Chihuahua), Working-group (500 images of Boxer and 500 images of Newfoundland). The training dataset contained a total of 5,000 images, this was further divided into 2 sets – training set with 3,500 images (∼70% of random data) and validation set with 1,500 images (∼30% of random data).

D. *Image Classification*

Image classification is the systematic arrangement of data into groups and categories based on similarities in features. For this work, the feature extraction and classification were done using our built convolutional neural network model. To train our built model using pytorch, we started by importing pytorch libraries, then we transformed the input images by resizing to 255 pixels, we centre cropped the images to 224 pixels which is our model's input size, we used totensor for converting the images into pytorch's usable format, afterwards we normalized the images and transformed the images. We loaded the images into the model via dataloader, then we moved the model into the Central Processing Device (CPU). With both the model and the training data defined, we configured the learning process by setting the training parameters for the model as learning rate of 0.001, batch size of 32, momentum of 0.9, epoch of 10, loss function as cross entropy loss and optimization function as Stochastic Gradient Descent (SGD). We set the model to training mode and initiated the training process. The validation dataset was used to evaluate the model's performance during the training process. After the completion of the training, we saved the model.

For the testing phase, we used our trained model to classify the test dataset into the respective dog groups. We started by setting the model to evaluation mode, we loaded the test dataset using a batch size of 1, afterwards we initiated the test. This work was validated by comparing its classification performance with two classic CNN models (ResNet-50 and SqueezeNet) using the same dataset, training parameters and testing parameters. This work was implemented using the Python programming language (Python 3.8.3) with pytorch open source software library on a laptop with the following configuration: Intel(R) Core ™ i3-2330 CPU@2.20GHz, 8GB RAM laptop running Microsoft Windows 10.

E. *Evaluation Metrics*

Evaluation metrics are indicators for assessing the performance of an experiment. In this work, accuracy, sensitivity, specificity and AUC were selected as the quantitative evaluation metrics. The accuracy, sensitivity and specificity were calculated using Eqns. (6), (7) and (8) respectively.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (7)$$

$$Specificity = \frac{TN}{TN+FP} \quad (8)$$

Receiver operating characteristic (ROC) curve is a probability curve that plots the True Positive Rate (TPR) on the y-axis and False Positive Rate (FPR) on the x-axis. The false positive rate and true positive rate can be represented using Eqns. (9) and (10).

$$False\ Positive\ Rate\ (FPR) = \frac{FP}{FP+TN} \quad (9)$$

$$True\ Positive\ Rate\ (TPR) = \frac{TP}{TP+FN} \quad (10)$$

Where TN, FN, FP TP and are true negative, false negative, false positive and true positive respectively. Area under curve (AUC) is defined as the area under the ROC curve and is an important evaluation indicator. It is given mathematically as shown in Eqn. (11).

$$AUC = \sum_{i\varepsilon\ positive\ class} \frac{rank_i - \frac{M(1+M)}{2}}{MXN} \quad (11)$$

where M is the number of positive class samples, N is the number of negative class samples, and $rank_i$ refers to positive sample score greater than negative samples. When AUC value is close to 1, the authenticity of the classifier is higher; when it is equal to 0.5, its authenticity is the lowest and it has no application value.

IV. RESULTS AND DISCUSSION

A. *Training result of the built CNN model*

For the proper operation of the built CNN model, it was trained using our training dataset. We trained the model for 10 epochs and it was observed that during the training process the training loss and validation loss were high at the start but decreased as the number of epochs increased. The model achieved training loss of 0.00007 and validation loss of 0.00009 at the 10th epoch. Figure 5 shows the training loss and validation loss plot. From Figure 5, it was observed that the training loss decreased rapidly in the first three epochs while the validation loss remained almost flat after the first epoch. The result indicated that the model generalized well on the validation dataset.

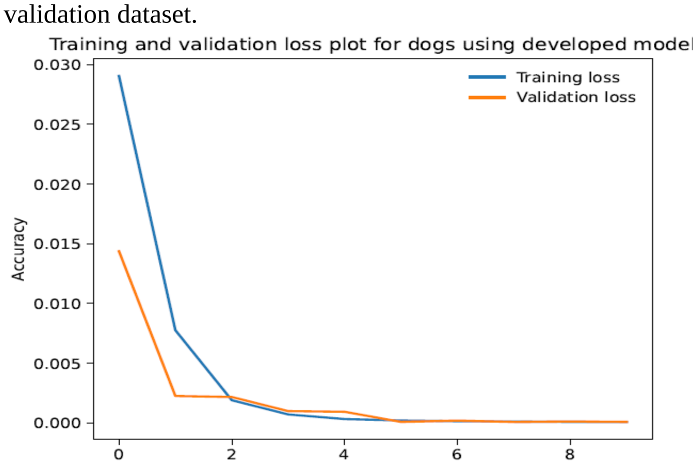

Figure 5: Training loss and validation loss plot

B. Testing result of the built CNN model

The performance of our built CNN model was tested using the test dataset and the result is as shown in Figures 6 (a) and 6(b).

In this research, a true positive occurs when the actual dog group is predicted correctly, any other prediction will be considered a false negative. Figure 6(a) shows the confusion matrix result obtained for the classification of dog breeds, the figure presents the predicted and actual class labels. The misclassification observed in the dog groups could be as a result of intra-breed differences and inter-breed similarities amongst dog breeds. The evaluation parameters were calculated using the values in the confusion matrix and Eqns. (8), (9) and (10). The result obtained is as shown in Table 2. In Figure 6(b), the result of the ROC analysis specified how well the positive class in each dog group was well separated from other groups, a higher true positive rate (TPR) and lower false positive rate (FPR) is desirable because the positive class is expected to be correctly classified.

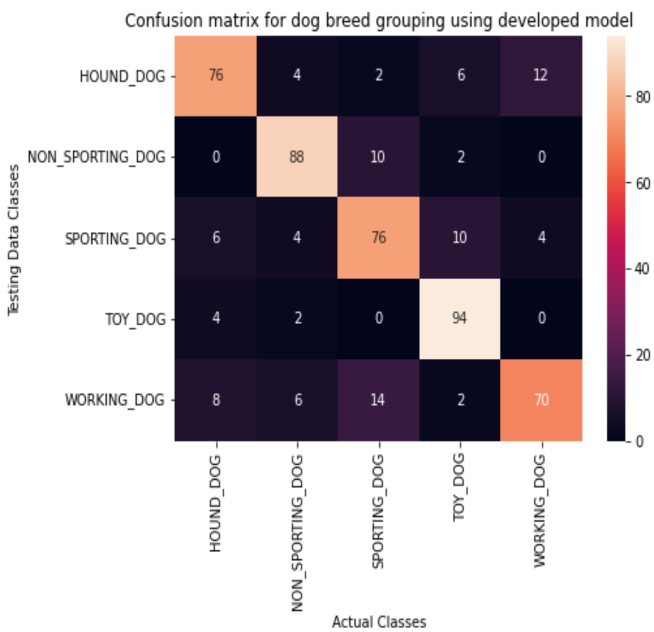

Figure 6 (a): Confusion matrix result

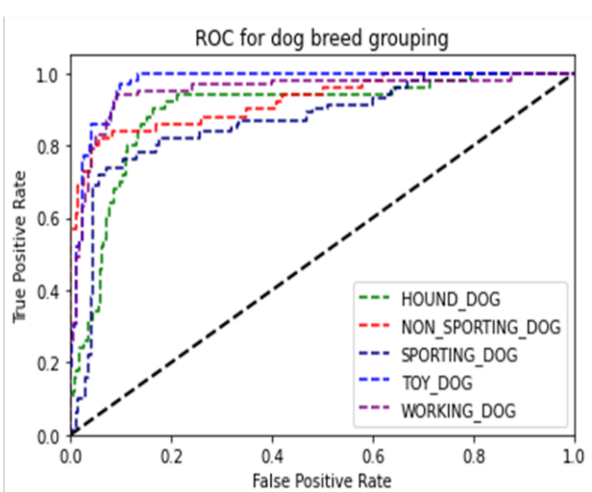

Figure 6(b): ROC plot result

From Table 2, All the dog groups had AUC value between 0.5 and 1, this indicated that the model was able to distinguish the negative class from the positive class as the model detected a greater number of true positive (TP) and true negative (TN) than false negative and false positive. Toy_dog had the highest accuracy, sensitivity and AUC of 95.0%, 95.0% and 97.0% respectively. This could be as a result of the differences in physical appearance between the dog breeds in toy_dog and dog breeds in other groups. The overall performance of the system was determined by calculating the average accuracy, average sensitivity, average specificity and average AUC as shown in Table 3.

Table 2: Performance of the developed model per dog group.

Parameters	Hound_Dog	Non_Sporting_Dog	Sporting_Dog	Toy_Dog	Working_Dog
Accuracy (%)	92.0	94.4	90.0	94.8	90.8
Sensitivity (%)	76.0	88.0	76.0	94.0	70.0
Specificity(%)	95.5	96.0	93.5	95.0	96.0
AUC (%)	95.0	93.0	87.0	97.0	95.0

Table 3: Overall performance of the developed technique.

Performance evaluation metric	Result (%)
Accuracy	92.4
Sensitivity	80.8
Specificity	95.2
AUC	93.4

From Table 3, it was observed that our technique achieved an overall accuracy of 92.4%, sensitivity of 80.8%, specificity of 95.2% and AUC of 93.4%. The work was able to establish that there were similarities in features amongst dog breeds. The model was 92.4% accurate. The overall AUC of 93.4% indicated that the model separated the five groups of dog breeds very well. 80.8% sensitivity indicates that the model correctly classified 80.8% of the positive samples. Hence, the developed model can be trusted in its ability to detect more positive samples.

C. Validating the performance of our proposed technique with classic CNN models

These section compares the result of our developed model with that of two classic models (ResNet-50 and SqueezeNet) in the classification of dog breeds into groups using the stanford's dog dataset with the same training parameters and testing parameters. The confusion matrix obtained is as shown in Figures 7(a) and 7(b). The validation result is provided in Table 4.

From Table 4, it was observed that all the three models performed well in the classification of dog breeds into groups. However, our proposed model had the least performances in terms of accuracy, sensitivity and specificity. This could be as a result of the fact that Resnet-50 and SqueezeNet were pretrained models, they were trained on large datasets. This enabled them to use the features and weights already learnt as a starting point. On the other hand, the proposed model was trained from scratch using few datasets. Hence, the performance of our proposed model could be improved by increasing the size of the dataset.

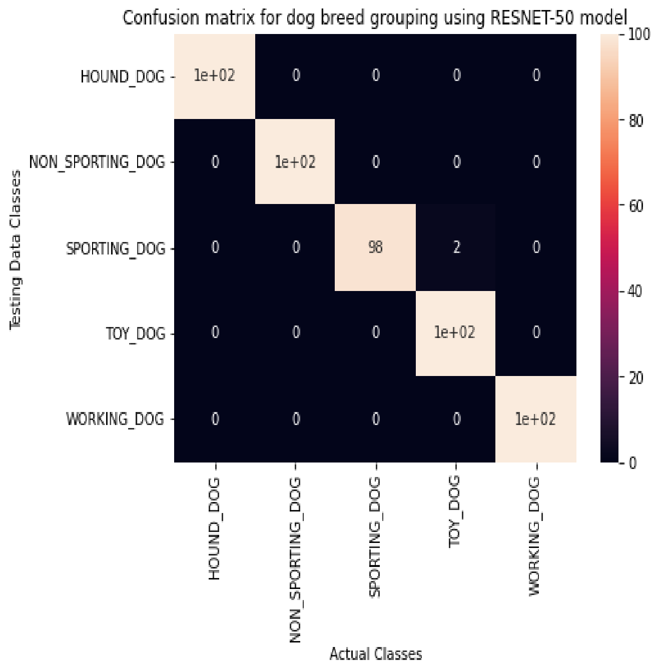

Figure 7(a): Confusion matrix for ResNet-50 model

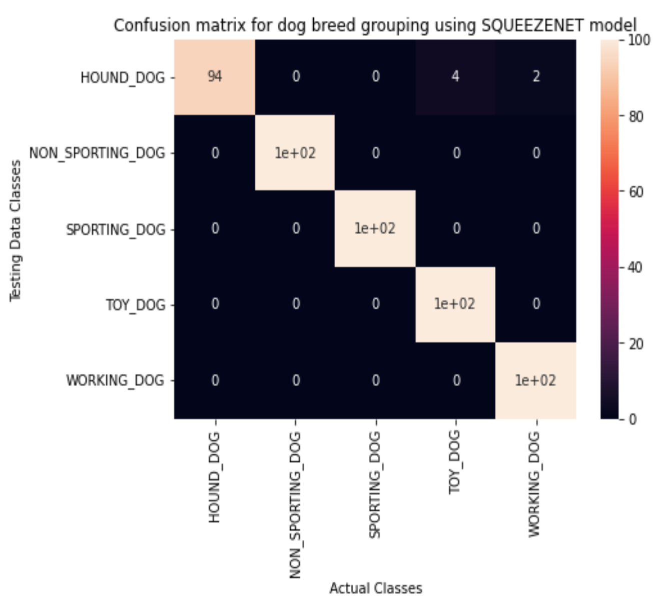

Figure 7(b): Confusion matrix for SqueezeNet model

Table 4: Validating the performance of our proposed model with classic CNN models

Model	Proposed model	SqueezeNet	ResNet-50
Accuracy	92.4	99.7	99.8
Sensitivity	80.8	98.8	99.6
Specificity	95.2	99.6	100.0
AUC	93.4	97.8	100.0

D. *Comparison of proposed technique with existing techniques*

Table 5 shows the comparison of our proposed technique with existing techniques that used the same dataset (Stanford's dog dataset) for dog breed classification. From Table 3, although existing works used only accuracy as their evaluation metric; our technique outperformed the work of Zalan *et al*, 2018, Durga *et al*, 2019, Kanika *et al*, 2020 and Akash *et al*, 2021 in terms of accuracy. This performance could be attributed to the built CNN model and the size of the dataset used.

Table 5: Comparison of the proposed technique with existing technique

Author (s)	Dataset	Technique	Classification Performance (%)
Zalan et al, (2018)	Stanford dog dataset	NASNet-A	Accuracy = 80.72
		Inception-Resnet-V2	Accuracy = 90.69
Durga et al, (2019)	Stanford dog dataset	Inception-V3	Accuracy = 89.0
		Inception-ResNet-V2	Accuracy = 94.0
		VGG 16	Accuracy = 81.0
		Xception	Accuracy = 93.0
Kanika et al, (2020)	Stanford dog dataset	Resnet 101	Accuracy = 71.63
		Resnet 50	Accuracy = 63.78
		InceptionResnet-V2	Accuracy = 40.72
		Inception-V3	Accuracy = 34.84
Akash et al, (2021)	Stanford dog dataset	Inception-V3	Accuracy = 85.0
		VGG-16	Accuracy = 69.0
Proposed technique	Stanford dog dataset	Built CNN model	Accuracy = 92.4 Sensitivity = 80.8 Specificity = 95.2 AUC = 93.4

V. CONCLUSION

In this work, dog breeds were automatically classified into groups based on the American Kennel Club (AKC) standard using a built convolutional neural network model. The research was performed using the open source Stanford's dog dataset and the findings indicated that the developed model performed excellently in terms of accuracy, sensitivity, specificity and AUC when compared to existing techniques that used the same dataset. Our proposed model's performance was validated with two classic CNN models (ResNet-50 and SqueezeNet) using the same datasets and parameters. This study has the following limitations; the study relied on convolutional neural network model for automatic feature extraction and classification, it might not capture the specific features of dog breed that are relevant for behaviuoral classifications. Hence, this work can only classify correctly the dog breeds it was trained with at high accuracy. Furthermore, the study uses a single dataset that may not represent the diversity and variability of dog breeds and their behaviour in real world scenario. Future work could merge the available datasets together or compare the performance of the developed model on different dog datasets. The study relied on the dog grouping experiment that was performed by the American Kennel Club standard; therefore, it did not consider other factors such as training, environment, health and personality that could influence dog behaviour. In future research, the number of dog breeds per class can be increased. Also, the number dog group can be increased to seven.

AUTHOR CONTRIBUTIONS

P.O. Adejumobi: Conceptualization, Design and Writing; **I.O. Adejumobi:** Data Collection and Writing; **O.A. Adebisi:** Methodology and Analysis; **S.O. Ayanlade:** Data Interpretation and Writing; **I.I. Adeaga:** Review and editing.

REFERENCES

Adebisi, O.A.; J.A. Ojo and O.T. Bello (2020). Computer aided diagnosis system for classification of abnormalities in thyroid nodules ultrasound images using deep learning. IOSR Journal of Computer Engineering (IOSR-JCE), 22(3): 60-66.

Aditya, K.; J. Nityananda; Y. Bangpeng and L. Fei-Fei (2013). Novel dataset for fine-grained image categorization: stanford dogs, pp. 1-2.

Akash, V.; K. Abhay; K. S. Aman and S.C. Surendra (2021). Dog breed classification using deep learning. International Conference on Intelligent Technologies (CONIT) Karnataka, India, 1-5, https://www.researchgate.net/publication/353693466.

Alex, K.; S. Ilya and E.H. Geoffrey (2012). ImageNet classification with deep convolutional neural networks. pp. 1-9, https://www.researchgate.net/publication/330729785.

Alexandre, B.L. and G.M. Mauricio (2017). Computer vision and artificial intelligence techniques applied to robot soccer. International Journal of Innovative Computing, Information and Control, 13(3):991-1005.

Althnian, A.; D. AlSaeed; H. Al-Baity; A. Samha; A.B. Dris; N. Alzakari; A. Abou Elwafa and H. Kurdi (2021). Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical. Applied Sciences, 11(796):1-18.

American Kennel Club (ACK) is the recognized and trusted expert in breed, https://www.akc.org.

Amit, K.J.; S. Mritynjay and K.S. Anjani (2021). Dog breed classification using convolutional neural network. International Journal of Swarm Intelligence, 6(2):130-142.

Anuj, G.; K. Rishi and P. Kailash (2019). A comparative study of fruit recognition system using machine learning approaches. International Journal of Engineering Technology and Applied Science (IJETAS), 5(7): 1-9.

Ayan, M.; S. Subhankar and J. Vinod (2022). A convolutional neural-network-based approach for automatic dog breed classification using modified-xception model. Electronic Systems and Intelligent Computing, Springer, 61-70.

Bickey, K.S.; K. Aman and K. Amrit (2020). Dog breed classifier for facial recognition using convolutional neural networks. 3rd International Conference on Smart Systems and Inventive Technology (ICSSIT), IEEE, 4(7):234-242.

Deepika, J.; V. Sowmya and K.P. Soman (2014). Image classification using neural network, *International Journal of Advancement in Research and Technology, 3(6):1661-1668.*

Dog Breed Classifier Using Deep Learning Jaypee University of Information Technology Waknaghat, Solan 173234, Himachal Pradesh July 2019.

Durga, C.; J.W. Lisa; R. Friederike; A. Nadja; S. Jessica and V. Zsofia (2019). Behavioural and cognitive changes in aged pet dogs: No effects of an enriched died and lifelong training. PLoS ONE, 15(9):34-43.

Enya, V. P.; H. Amanda; J. Per and S. V. R. Lina, (2022). Breed group differences in the unsolvable problem task: herding dogs prefer their owner, while solitary hunting dogs seek stranger proximity. Animal Cognition, 25:597–603.

Ezat, W. A.; M.M. Dessouky and N.A. Ismail (2019). Multi-class image classification using deep learning algorithm. The 4th International Conference on Advanced Technology and Applied Sciences (ICaTAS), 167-172, https://www.researchgate.net/publication/335821715.

Federation Cynologique Internationale (FCI), an international federation of a number of national kennel Club www.fci.be.

Ghirlanda, S.; A. Acerbi; H. Herzog and J.A. Serpell (2013). Fashion vs. function in cultural evolution: The case of dog breed popularity, PLoS ONE, 8(9): 747-770.

James, A. S. and L.D. Deborah (2014). Dog breeds and their behavior. 31-57, https://www.researchgate.net/publication/271826897.

Jiongxin, L.; K. Angjoo, K.; J. David and B. Peter (2012). Dog breed classification using part localization. 1-14, https://www.researchgate.net/publication/323456726.

Kanika, G.; G. Sharankhala; A. Bhavya and G. Riya (2020). Dog breed identification using CNN architecture. *International Journal of Scientific and Engineering Research, 11(6):1040-1046.*

Kenneth, L.; T. Xinyuan and Y. Svetlana (2019). Dog identification using soft biometrics and neural networks. International Joint Conference on Neural Networks (IJCNN), 5(7):1-7.

Kenth, S. and F. Bjorn (2012). Personality traits in the domestic dog (*Canis Familiaris*). Applied Animal Behaviour Science, 79(2): 234-246.

Manoj, M.K.; M. Neelima; M. Harshali and M.G.R. Venu (2018). Image classification using deep learning, *International Journal of Engineering and Technology, 7(2.7): 614-617.*

Maria, S. (2015). Dog breeds of the world. 1-96, http://www.plexidors.com

Middi, V.; R. Sai and A.H. Tanvir (2018). Dog Breed Classifier using Convolution Neural Networks. in IEEE., 6(2):1-13.

Mohamed, S.M.; S. Naveen; K.C. Praveen and M.R.M. Arun (2020). Dog breed identification using convolution neural network and web scraping. International Journal of Science and Research (IJSR), 9(3): 488-491.

Muhammad, A.; R. Ziaur; A.A. Waheed and T. Muhammad (2019). An Optimized Architecture of Image Classification Using Convolutional Neural Network. International Journal of Image, Graphics and Signal Processing, Vol. 10, 30-39.

Naufal, H. P.; R. Ema and K. Gamma, (2022). Classification of dog breeds from sporting groups using convolutional neural network. Jurnal Ilmiah Penelitian dan Pembelajaran Informatika (JIPI), 7(4):1080 – 1087.

Neda, A.; I. Tehran; T. Omid and A. Mohammad (2021). Flower image classification using deep convolutional neural network. https://www.researchgate.net/publication/352093979, pp. 1-6.

Neeraj, K.; K.B. Dashutosh; K.D. Rakesh and B. Rajendra (2019). Performance analysis of support vector machine in defective and non-defective mangoes classification. International Journal of Engineering and Advanced Technology (IJEAT), 8(4):1563-1572.

Philip, O.A; A. A. Oluwadare; I. A. Iyabo; B. B. Abiodun and M. A. Kolawole (2023). Automatic classification of mechanical components of engines using deep learning techniques. International Journal of Innovative Science and Research Technology (IJISRT), 8(4):2028-2037.

Punyanuch, B.; K. Worapan and T. Kittikhum (2021). Knowing your dog breed: identifying a dog breed with deep learning. International Journal of Automation and Computing, 18(1), 45-54.

Punyanuch, B.; T. Kittikhum and K. Worapan (2019). Breakthrough Conventional Based Approach for Dog Breed Classification using CNN with Transfer Learning. 11th Learning Information Conference on Information Technology and Electrical Engineering (ICITEE), 6(5): 2032-2043.

Ravi, K., and K. Shailender (2019). Image Segmentation Using Convolutional Neural Network. International Journal of Scientific and Technology Research, 8(11): 667-675.

Richard, O. S.; W. Fang and C. Wenbin (2018). A mobile application for dog breed detection and recognition based on deep learning', 2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT), 87-96.

Rikiya, Y., N. Mizuho, K.G.D. Richard and T. Kaori (2018). Convolutional neural networks: an overview and application in radiology, springer, 611-629, https://doi.org/10.1007/s13244-018-0639-9.

Sandra, V. and S. Remya (2021). Dog breed classification using CNN. 195-205, DOI:10.1002/9781119776529.ch.10.

Seyyid, A.M. (2015). A comparative study of feature extraction methods in image classification. International Journal of Image, Graphics and Signal Processing, 3:16-23.

Sneha, I. K.; S.K. Shubhada and G.K. Sharada (2020). Dog breed prediction using convolutional neural network. International Journal of Engineering and Advanced Technology (IJEAT), 9(5):318-322.

Suyash, B.; P. Rishikesh; W. Rohit; J. Kaustubh and B. Balaji (2021). Dog breed classification using convolutional neural network. International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 6(1):753-759.

Thai, H.L; T.S. Hai and T.T. Nguyen (2012). Image classification using support vector machine and artificial neural network. International Journal of Information Technology and Computer Science, 5(2):32-38.

Vikas, K.M.; K. Shobhit and S. Neeraj (2017). Image Acquisition and Techniques to Perform Image Acquisition. SAMRIDDHI-A Journal of Physical Sciences, Engineering and Technology (S-JPSET), 9(1):21-24.

Vikram, K. and R.N. Awale (2020). A facial EMG data analysis for emotion classification based on spectral kurtogram and CNN. International Journal of digital signals and smart systems, 4(1/2/3):50-63.

Wenting, S.; C. Jiaquan; L. Muyun and L. Fangyu (2015). Dog breed identification. 1-6, https://noiselab.ucsd.edu.

Wenzhong, L. (2020). Interfruit: Deep Learning Network for Classifying Fruit Images. 1-16, https://doi.org/10.1101/2020.02.09.941039.

Whitney, L.; M. Brian and S. Vijay (2015). Dog breed identification. 1-7, *https://web.stanford.edu.*

Xiaolong, W.; L. Vincent; S. Scott and K. Chandra (2015). Dog breed classification via landmarks. 5237-5241, https://www.researchgate.net/publication/283183525.

Xinyuan, T; L. Kenneth and Y. Svetlana (2018). Transfer learning on convolutional neural networks for dog identification. IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), 9(3):123-132.

Zalan, R.; S. Csaba; V. Zsolt and Z. Attila (2018). Dog breed identification using deep learning. 1-6, https://www.researchgate.net/publication/328834665.