Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/6/20230968

Automatic dog breed classification using deep learning

T Kumaravel[1, 3], P Natesan[1], V E Sathishkumar[2], Sathiya Shri N[1], Swathy G[1], Uvetha V[1]

[1]Department of Computer Science andEngineering, Kongu Engineering College, Erode, Tamil Nadu, India
[2]Department of Software Engineering, Jeonbuk National University, Jeonji-si, Jeollabuk-do, Jeonju-si, Republic of Korea

sathish@jbnu.ac.kr

Abstract. Dogs are a common type of animal that can present various problems, such as issues with population control, controlling rabies outbreaks, administering vaccinations, and legal ownership, due to their large numbers. Understanding the breed of the dog can help the owner identify potential health issues and determine their lifespan. There are over 120 different breeds of dogs, each with unique characteristics and health concerns. To properly care for and train a dog, it is important to know their breed. This study discusses methods of classifying dog breeds and presents a method using a CNN (Convolutional Neural Network) to accurately identify different breeds by analyzing dog images. This approach, using DenseNet201, achieved an accuracy of 87.34% on the Dog Breed Images dataset and is more effective than other methods found in literature.

Keywords: Dog breed, Convolutional Neural Network (CNN), Dog Breed Classification Dataset, DenseNet, Xception, Inception

1. Introduction

Image classification has a wide range of applications beyond human facial recognition in the digital world, making it an increasingly significant field of study. There are various essential uses for image classification, particularly in identifying specific elements in pictures, in fields such as social media, marketing, and national security [1]. In addition, the use of dog breed identification models can be used to predict behavior, match dogs with their owners, locate lost dogs and support targeted advertising. The rise of social media in the last decade has created many opportunities for people to share pictures of their pets, which can lead to potential for tailored marketing [2].

A methodology for identifying dog breeds in images can aid businesses in determining which breeds are popular and incorporating them into their marketing strategies. A dog breed classification model is constructed utilizing the Stanford Dog Dataset and a variety of transfer learning methods, as well as from scratch models for 120 different breeds [3]. The excessive number of dogs leads to numerous problems, including social control, reducing rabies epidemics, vaccination control, and legal ownership. Therefore, it is crucial to identify the breeds of dogs in order to provide appropriate therapies and training. The research presented here includes the classification of 120 categories of dog

DOI: 10.54254/2755-2721/6/20230968

breeds. Fig. 1 illustrates the images used in the proposed work, utilizing Inception V3, MobileNetV2, DenseNet201, and Xception.

2. Related work

It has been observed that various deep learning models that utilize reputable deep learning architectures are utilized for the classification of dog breeds. Furthermore, many researchers have published recent versions to enhance the accuracy of dog breed classification in a range of circumstances. Recently, for classification of dog breeds, deep convolutional neural networks have been successfully applied, resulting in better performance on prediction tasks. A paper titled "Examining Conventional Breed Group Classification's Usefulness as a Justification for Domestic Dog Issue Resolution" was proposed by Daniel Mills, Jonathan Cooper, and Tracey Clarke where, SPSS 14 was used to examine the significant difference in performance. A paper on "Dog Breed Classification Using CNN" was proposed by Remya and Sandra Varghese, in which the ResNet-50 model of transfer learning is used for classifying the breed lesion-related layer regions to improve classification. Additionally, a paper on "On-Farm Automatic Sheep Breed Classification Using Deep Learning" was proposed by Sanabel Abu Jwade and Ajmal Mian, in which various features of sheep such as shoulder height, weight, and color are collected successfully from the dataset and classified using VGG-16. The studies demonstrate that Deep learning architectures are progressively being used in the classification of dog breeds from images. However, there are still a number of issues that need to be resolved before Deep learning architectures can be effectively used, such as shorter training times and fewer parameters [4].

3. Materials and methods

3.1. Datasets

In the early stages of the project, data collection is a crucial component. Images of various breeds were collected from the Stanford Dataset and Kaggle [5]. The images that were downloaded from Kaggle were pre-processed. This dataset includes multiple classes and a variety of images. Image augmentation techniques such as horizontal flipping, padding, cropping, and rotation were performed on the existing training data to create unique and individualized training images and to reduce overfitting of the model. After augmentation, the dataset consisted of around 120,000 images [6]. A unique dataset was generated from the neutralized data gathered from the Stanford dataset, which acted as a public dataset on Kaggle named "Dog Breed Classification." After selecting the dataset, the training dataset and testing dataset were divided by allocating 10% for testing, 10% for validation, and 80% for training the model, which can be used for classification. This was done to ensure that there is ample data available for training, resulting in a precise model. In total, 1000 images were used in each class, out of which 100 images each were used for testing and validation and 800 images for training.

3.2. Convolutional neural networks

The convolutional neural network (CNN) is a widely recognized and popular deep learning model. It is a type of deep learning algorithm that is utilized for various challenges involving image classification and is gaining recognition in industries such as music and health. The CNN comprises various layers, including convolutional layers, pooling layers, and fully connected layers, which are used to dynamically learn data topologies using the back-propagation technique [6].

3.2.1. The Convolutional layer. Convolution is the first and most important layer of CNN. Simply applying a filter to an input results in a convolution, which shrinks the size of the image while simultaneously condensing all of the field data as a only one pixel. In Fig. 2 an image convolution that recognizes the creature's edges is displayed.

3.2.2. The pooling layer. Max pooling is utilized as a method for reducing the input image size for convolutional layers. It is commonly known that pooling layers, including maximum, global, and average pooling, are used in neural networks to decrease the number of features that need to be processed and the computational requirements of the network. The application of a max pooling layer results in a reduction of the image dimensions as depicted in Fig. 3[5].

3.2.3. Activation function. The activation function used to determine the decision of whether to stimulate a neuron is the Rectified Linear Unit (ReLU) before the information is passed on to the next layer of neurons. Activation functions, such as ReLU, serve as a non-linear adjustment that allows the neural network to learn more complex representations of the data [7].

4. Proposed models

It is the identification and categorization of dogs and their breeds is crucial for providing proper care and training [8]. While a number of standard approaches, both simple and complicated, can be used for this purpose, they often require significant time and human labor. To improve the usefulness of dog breed classification in various contexts, some researchers have incorporated modified variants. The proposed system uses dog images as input and employs a variety of deep learning models to test the effectiveness of pre-processing techniques such as equalization, enhancement, augmentation, and combination. The pre-processed data is then fed into multiple models with various parameters in order to select the optimal model for classifying dogs according to their breeds. The system for identifying dog breeds is illustrated in Fig. 1.

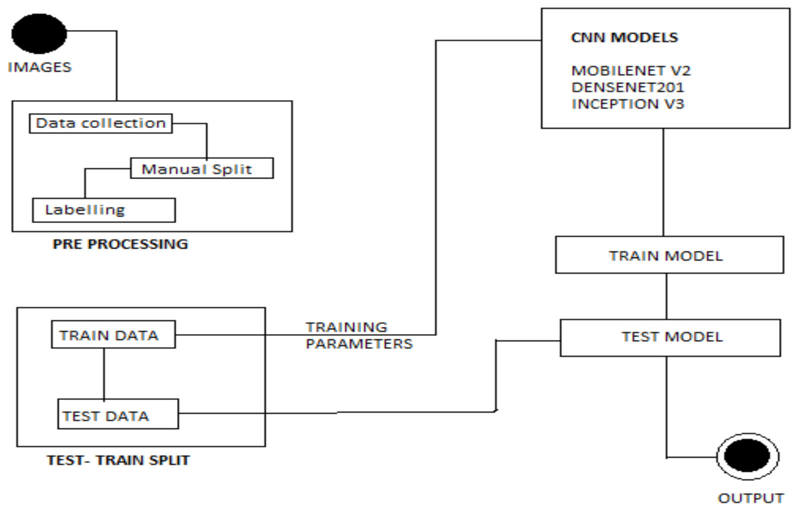

Figure1. Proposed workflow.

4.1.1. MobileNetV2. A CNN network design known as the MobileNetV2 model functions almost flawlessly on mobile devices. In order to filter features, lightweight depthwise convolutions are utilised as a reason for the intermediate expansion layer's nonlinearity. The MobileNetV2 design has 19 extra bottleneck layers in addition to the 32-filter initial fully convolution layer. Figure 2 shows MOBILENETV2 model's basic design[9].

DOI: 10.54254/2755-2721/6/20230968

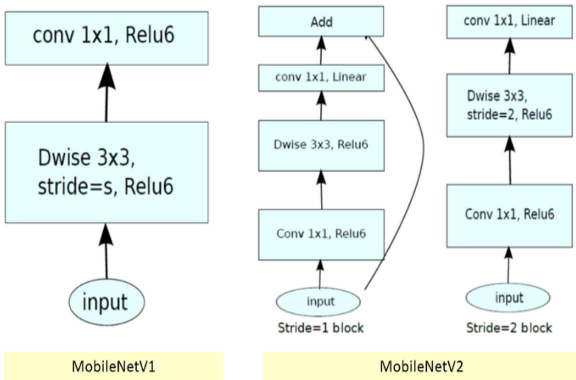

Figure 2. Architecture of MOBILENETV2.

4.1.2. Densenet201. DenseNet201 is one of the type of CNN network that often utilizes Dense Blocks to build denser connections in between the layers so that every level is directly connected (with matching feature-map sizes). Each of the layer is directly associated to each and every layer in a feed-forward strategy (within each dense block). The general design of the DenseNet201 model has been portrayed in Fig.3[5].

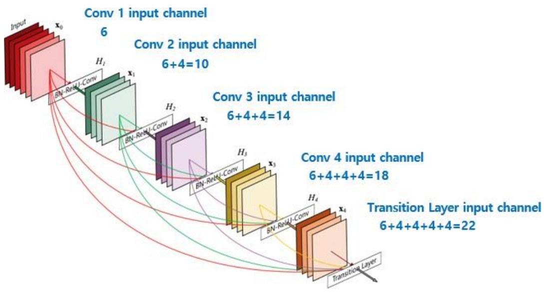

Figure 3. Architecture of DenseNet201.

4.1.3. Inception V3. Inception-v3 is kind of CNN network design, among other advancements, makes use of Label Smoothing, Factorized 7 x 7 convolutions, and the usage of an auxiliary classifier to convey label information further down the network (along with the use of batch normalisation for layers in the side head). The Inception model's general design is seen in Fig. 4[5].

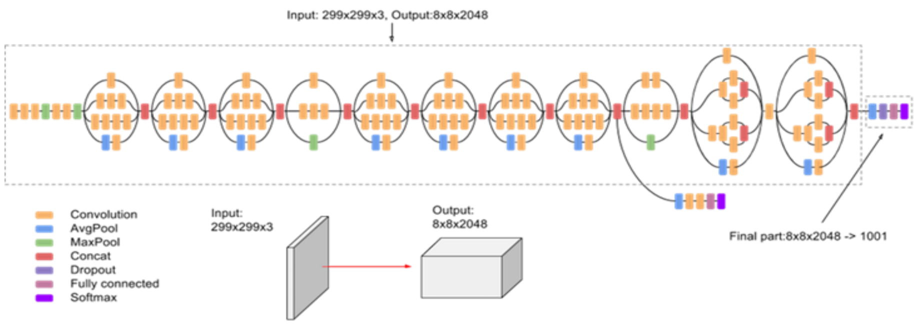

Figure 4. Architecture of InceptionV3.

5. Result and discussion

The MobileNetV2, InceptionV3, Xception and DenseNet201 models are used in the study, results are given and analyzed in this section. For the experimental scenarios, all images are resized to 224* 224* 3 pixels to allow for input of DenseNet, Xception and Inception models. The accuracy and loss

DOI: 10.54254/2755-2721/6/20230968

for validation and training of models, which run the varied numbers of epochs, are displayed below [10].

Densenet201 model results are presented in Table 1 for various epoch counts. One epoch was used for training and testing the Densenet201 model initially. Then, as precision rose along with the number of epochs, the number of epochs was gradually raised. Densenet201 model reaches training accuracy of 87.34% across 25 epochs. But there has been no appreciable increase in accuracy after 20 epochs. The lack of sufficient photos is the reason for this.

Table 1. Performance of Densent201.

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.7000	2.2832	0.7931	1.7886
5	0.9542	0.3999	0.8425	2.4255
10	0.9789	0.2176	0.8684	2.7704
15	0.9851	0.1653	0.8658	3.4832
20	0.9892	0.1311	0.8742	3.6185
25	0.9917	0.1048	0.8738	4.0261

Table 2 displays how well the Xception model performed. The correctness is verified, and the number of epochs is steadily raised, just like in Xception. It is set to learn at a rate of 0.0001. For 25 epochs, training accuracy was 85.75%.

Table 2. XCEPTION performance.

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.7119	2.7403	0.7566	2.7671
5	0.9370	0.5603	0.8203	3.1545
10	0.9682	0.3102	0.8303	3.3231
15	0.9795	0.2169	0.8556	0.8354
20	0.9850	0.1668	0.8511	4.1961
25	0.9882	0.1346	0.8643	4.5408

In Table 3, the MobileNetV2 model's performance is displayed. It is set to learn at a rate of 0.0001. 82.99% of the training accuracy was attained throughout 25 epochs.

Table 3. MOBILENETV2 performance.

EPOCH	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.6832	3.3092	0.7532	3.1391
5	0.9320	0.6369	0.7940	3.8291
10	0.9659	0.3623	0.8082	4.8555
15	0.9785	0.2439	0.8012	6.3592
20	0.9831	0.2093	0.8102	6.7461
25	0.9836	0.5160	0.8114	7.5588

In Table 4, the InceptionV3 model's performance is displayed. It is set to learn at a rate of 0.001. 80.90% of the training accuracy was attained throughout 25 epochs.

DOI: 10.54254/2755-2721/6/20230968

Table 4. Performance of INCEPTIONV3 (continue).

EPOCH	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.6832	3.3092	0.7532	3.1391
5	0.9320	0.6369	0.7940	3.8291
10	0.9659	0.3623	0.8082	4.8555
15	0.9785	0.2439	0.8012	6.3592
20	0.9831	0.2093	0.8102	6.7461
25	0.9836	0.5160	0.8114	7.5588

The confusion matrix obtained for the DenseNet201 architecture is portrayed in Fig 5. Accurate classifications are presented by the confusion matrix's diagonal entries. On contrary, misclassifications are the remainder. X axis shows expected classes, while Y axis shows actual classes. Confusion matrix is used to calculate individual class measures such as precision, F1 score ,accuracy and recall. The metrics values are determined by using indices like FP, FN, TP and TN.

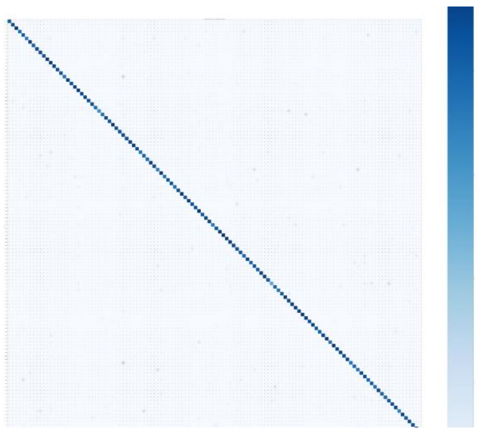

Figure 5. Confusion matrix for DenseNet201 model.

6. Conclusion

The ability to tell whether one's dog is prone to any breeding-related issues depends on knowing what breed it is. Knowing this information can help the owner predict how long their dog will live and how to best suit his needs. The breed can also have an impact on lifespan. The most recent developments have made it simpler and more efficient to get a required outcomes in deep learning technology. DenseNet201 is used most frequently because it considerably reduces the number of parameters, enhances feature propagation, encourages feature reuse, and resolves the vanishing-gradient problem. Different models and implementation strategies have been finished and put to rest. The proposed method, which employs DenseNet201, is demonstrated to have a high recognition rate and provide a average accuracy of about 87.34%. It can be used for categorizing dog breeds in detail.

References

[1] Sandra Varghese and Remya S. "Dog breed classification using CNN", pp. 1097-1105. 2021.

[2] Sanabel Abu Jwade and Ajmal Mian, Andrew Guzzomi, "On farm automatic sheep breed classification using deep learning," Animals, vol. 2, no. 2, pp. 301–315, 2019.

[3] Tracey Clarke, Daniel Mills and Jonathan Cooper, "Exploring the utility of traditional breed group classification as an explanation of problem solving behaviour of domestic dogs, vol. 7, no. 5, pp. 470–482, 2019.

DOI: 10.54254/2755-2721/6/20230968

[4] R. Kumar M. Sharma K. Dhawale and G. Singal "Identification of Dog Breeds Using Deep Learning" Proc. 2019 IEEE 9th Int. Conf. Adv. Comput. IACC 2019 pp. 193-198 2019.
[5] Malliga Subramanian, Kogilavani Shanmugavadivel, Obuli Sai Naren, K Premkumar, K Rankish. "Classification of Retinal OCT Images Using Deep Learning", 2022 International Conference on Computer Communication and Informatics (ICCCI), 2022
[6] KAGGLE – Dog Breed Clasification Images https://www.kaggle.com/datasets/uvetha/dog-breeds-rrr
[7] Rajalaxmi, R. R., Saradha, M., Fathima, S. K., Sathish Kumar, V. E., Sandeep kumar, M., & Prabhu, J. (2022). An Improved MangoNet Architecture Using Harris Hawks Optimization for Fruit Classification with Uncertainty Estimation. Journal of Uncertain Systems.
[8] Subramanian, M., Kumar, M. S., Sathishkumar, V. E., Prabhu, J., Karthick, A., Ganesh, S. S., & Meem, M. A. (2022). Diagnosis of retinal diseases based on Bayesian optimization deep learning network using optical coherence tomography images. Computational Intelligence and Neuroscience, 2022.
[9] Subramanian, M., Rajasekar, V., VE, S., Shanmugavadivel, K., & Nandhini, P. S. (2022). Effectiveness of Decentralized Federated Learning Algorithms in Healthcare: A Case Study on Cancer Classification. Electronics, 11(24), 4117.
[10] Shanmugavadivel, K., Sathishkumar, V. E., Kumar, M. S., Maheshwari, V., Prabhu, J., & Allayear, S. M. (2022). Investigation of Applying Machine Learning and Hyperparameter Tuned Deep Learning Approaches for Arrhythmia Detection in ECG Images. Computational & Mathematical Methods in Medicine.