Dog Breed Classification Using Deep Learning

Conference Paper · June 2021

DOI: 10.1109/CONIT51480.2021.9498338

4 authors, including:

Akash Varshney

1 PUBLICATION **7** CITATIONS

SEE PROFILE

Aman Kumar Singh

Galgotias University

2 PUBLICATIONS **10** CITATIONS

SEE PROFILE

Surendra singh Chauhan

Chandigarh University

22 PUBLICATIONS **58** CITATIONS

SEE PROFILE

2021 International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 25-27, 2021

Dog Breed Classification Using Deep Learning

Akash Varshney
Computer Science and Engineering
Galgotias University
Greater Noida
varshneyakash51@gmail.com

Abhay Katiyar
Computer Science and Engineering
Galgotias University
Greater Noida
abhaykatiyar212@gmail.com

Aman Kumar Singh
Computer Science and Engineering
Galgotias University
Greater Noida
thakuramansingh1999@gmail.com

Surendra Singh Chauhan
Computer Science and Engineering
Galgotias University
Greater Noida
surendra.chauhan@galgotiasuniversity.edu.in

***Abstract*— In this paper, we have proposed two models to classify dogs according to their breeds. Since the classification of dogs is becoming very difficult and moreover, these classifications are taken on the deep learning concept and training a fully defined data set helps in training both models which predicts the different accuracy levels at both ends. Since every now and then predictions are taken for every model. During the study, we came across many types of functionality levels that were not taken in previous studies too. Also, our approach also works on the main concept of transfer learning which deals with data augmentation technique with its properties to increase the size of data set, after which accuracy levels are matched or it is compared with both the models so that a comparison can be made for both the models and the classification is also done with a profound approach.**

Keywords*— *Deep Leaning, Machine Learning, Dog Breed, Classification

I. INTRODUCTION

Dog has been a companion for this world and for human mankind for the last many years. But now looking around and seeing dogs we find that there are many types of dogs who look some way or the other way different from each other[1]. Also, if we carry forward this, we found that there are many different kinds of dogs which are looking a bit similar but are of different breeds. These dogs are not only different in their breeds but they also differ in many of their characteristics such as their behavior towards humans, liking and disliking habits of each other, etc. Not only this one of the reasons for the classification of dogs according to their breed is that now in today's present time breeding of a dog with a different kind of dog has become very usual due to which every now and then a new variety of dog breeds are found. Since breeding mainly depends upon the matting of two species and on the other hand, it also depends upon the inner genetics of mother and father due to which matting is done and new breeds are produced which is not able to identify.

Recently dogs can be classify using an expert-based approach. These are one who has a variety of knowledge of different breeds of dogs. Hopefully, this is very difficult as experts are not available, and secondly, there is a DNA approach. But this DNA approach is expensive and time taking as at present time there are a total of 20,580 dogs breed present in the world.

In this paper, we are using an effective approach which is known as transfer learning. In this method, we have used pre-trained models which are used to classify the dog breeds. As a whole, we have worked on two neural networks which are Inception V3 and VGG16 which are trained on the Stanford Dog dataset and their accuracy is compared.

Fig. 1. Example of dog breeds

In the continuation of this research, we have tried to evolve as such by making an increase in the accuracy and more of such as: -

1. As compared to all other approaches we are using the transfer learning approach using two neural networks which are VGG16 and Inception V3,

2. We have provided an increase in accuracy level of the previous model by changing the number of training steps i.e., number of epochs and changing batch size in the image data generator class where data augmentation is done changing the batch size has increased the levels in dataset moreover, using rotation_range, width_shift_range, height_shift_range, and all the other properties of data augmentation[2].

3. It provides the enhancement in the accuracy level. On the other hand, both the model has been taken with the same parameters and trained on the same dataset (Stanford dog dataset) which helped in comparing their accuracy level of both the models. Not only this but while comparing both the dataset it was clear that the functional criteria of both the models were different.

978-1-7281-8583-5/21/$31.00 ©2021 IEEE

So as this paper begins with the introduction to the dog breed classification. It describes all the important aspects and the introductory part of the paper. Second section continues with the related work means all the work which has been done in the past with their pros and cons. Moreover, it consists of all the methods which has been implemented and it also consists of our proposed model with its basic structure. Third section consists of the data set used in the model for its training purpose and for the validation use also. Fourth section belongs to the data augmentation where the augmentation is done using its all properties to make the dataset fit for training purpose [2]. Fifth and sixth section proceeds with the building of the both models used in the study i.e., VGG16 and Inception V3. Seventh section deals with the result taken out in the both models with their accuracy graph and for accuracy vs epoch and loss vs epoch.

II. RELATED WORK

From the study it has been found that many approaches have been tried or classification of dogs. One of the approach is using Local Binary Pattern(LBP) technique[1][3] which is used as a face recognition. This algorithm mainly provides labels to every pixel of an image using binary numbers. Not only this Histogram Oriented Gradient also known as HOG algorithm has also been used by making histograms based on distribution of directions of gradients[1].

PCA based method is also approached on this classification by coarse classifier and the fine classifier respectively[4]. Fine-grained classification describes its importance by differentiating a class from a subclass. Many subclasses are formed the classes which are used to classification[5]. Some of the classification techniques used in literatures were lacking behind in one or other form. Using direct Convolutional neural network[5][6] in its basic form is also suitable as not every algorithm or technique applied is perfect but taking out maximum result from the good approach is important.

In the past studies a most common approach i.e., using convolutional neural network as a basic model. In that model a feature extraction approach is carried out with the LBP images[1] these works as histograms of images and therefore these histograms are carried out using feature extraction method.

Since transfer learning is working best for the classification and also on the other hand accuracy is the main reason for the classification models. So transfer learning is used with the two pre-trained neural networks which are used to classify dogs according to their breeds[1]. In Fig. 2 and also in our study we have first carried out the data preprocessing of our data set which brings out every detail of our data set and moreover, it is best to have every information about the data set before we start building models for training the data set. After preprocessing, we carry forward with data augmentation where the data set is provided with datagenerator class with its attributes. After this both models are taken with transfer learning and the models are trained using the same data set and the accuracy is observed.

As the above study was done there comes out to be some pros and cons of some of the researches. The study made by Hsu, David[5] it deals with the CNN neural networks which work on the architecture of CNN neural networks the accuracy which they bring about was 83 which was up to their mark, but the CPU machine which they have taken in their study takes up to 2 days to train about 1000 iterations, and the GPU machine taking about 8 hours to run 100,000 iterations.

Along with this study there comes another study that works on Grassmann manifold technique[7]. They have used landmark labels in their dataset and they also proved the performance of their model by keeping the same training and testing divisions.

The next study we come through was with the dog breed classification using part localization by Jiongxin Liu[2]. They work on a different technique that accurately localizes the eyes and the structure of the dog and manages to provide a difference in their structure. Their classification algorithm mainly focuses entirely on the face of the dog breed. But their study found some problems that some of the breeds of dogs were named different or the experimental phase of some dogs were different.

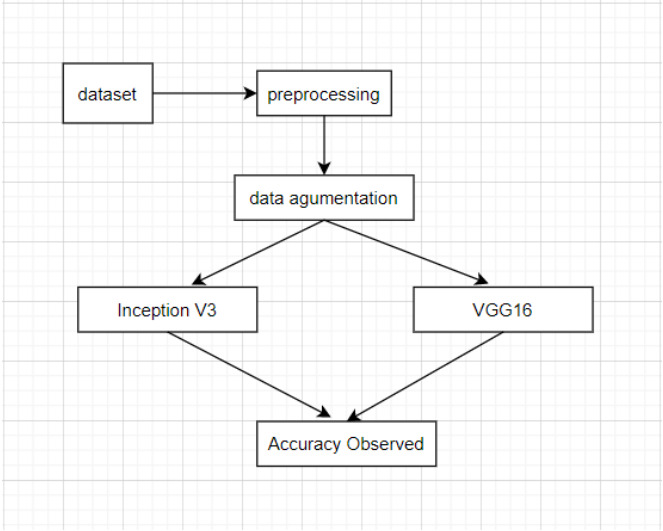

Fig. 2. Basic structure of proposed model

III. DATA SET

Dataset use for the classification model is Stanford Dog dataset which has all types of dog breeds around the world. This dataset contains images, annotations and the excel file which contains all the dog breed types. Detailing of this data set contains:

- Number of images: 20,50
- Number of categories:120
- Annotations: Class labels, Bounding boxes

This data set is very good as it holds mainly every type of image of dogs which provides a high level of accuracy level for training the data set and it also consists of 120 categories of dogs allowing us to make it a fully prepared dataset for our model. While working on this dataset we came across many of the different aspects of dogs which we have seen or not.

Data set is approx. 750MB in size with approx. 41.2k files in it.

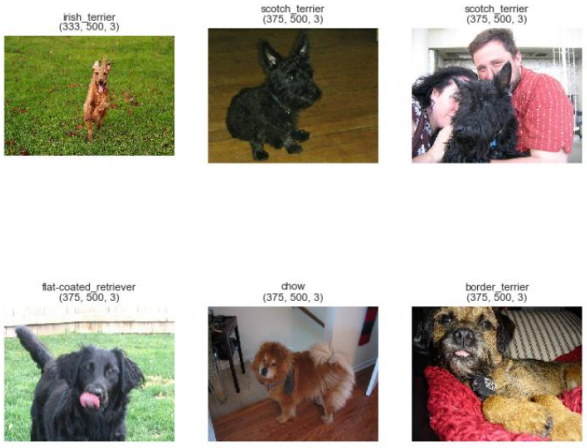

Fig. 3. Data set images used in training models

Before building the model, dataset is also preprocessed. With the help of preprocessing, we are able to get all the information in the dataset. It helped us to know the quality of data and on the other hand, it also helped to prevent data from redundancy. Preprocessing data set also provides an enhancement in both of our models which also plays an important role in our study.

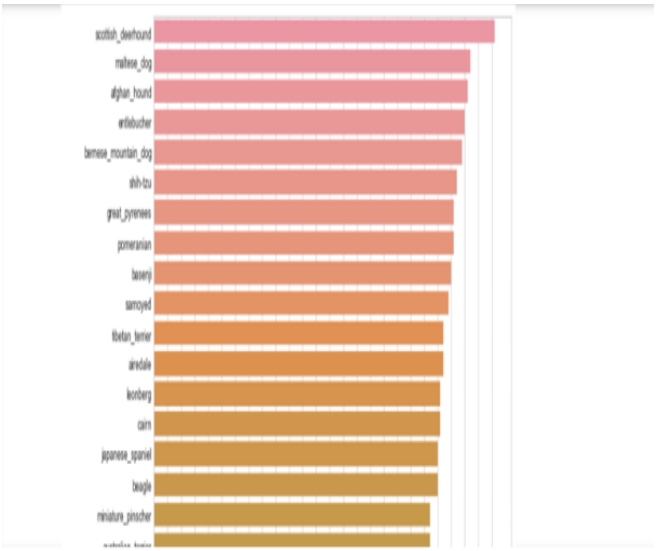

Fig. 4. Data Preprocessing

IV. DATA AUGMENTATION

The data analysis technique is applied which is used to increase the amount of data by adding an average of modified copies of already existing dataset or by newly originated data set [2]. Mainly the purpose of data augmentation is to remove overfitting and increasing he dataset for the training set. This technique has been implemented and the result is in Fig. 4, we can see that images are augmented according to the factors which we applying in the ImagedataGenerator class. These behaves like attributes to this class and helps in making new dataset.

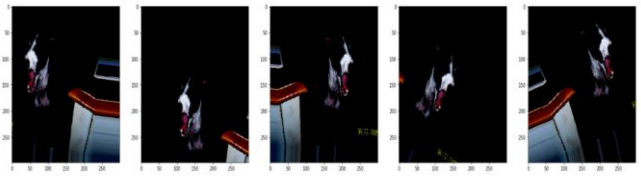

Fig. 5. Data Augmentation

V. INCEPTION V3

Inception V3[8] majorly known as image classifier neural network is the convolutional neural network and the third version of the inception model which was first developed by the GoogLeNet model. Since then, it has become a vast neural network and attained a good accuracy level in accordance to image classification. Some of the characteristics of this model includes weights, include_top, input_shape, and many more but if we talk about classification these are some of the most important characteristics for this model Inception v3 is an effective neural network as it helps to learn many of the complex and difficult features of a model or a data set. On the other hand, inception v3 has many filter sizes such as 1*1, 3*3, 5*5 which can be applied at the same level. done and new breeds are produced which is not able to identify [8].

Moreover, for a pretrained model, the transfer learning[9] is carrying out the weights for another model so that the accuracy level for the model is increased. In our model, the last layers of inception v3 are removed and the deep layers are added ending with a dense layer where the weights are taken as of ImageNet and the activation function is relu, including to which the dense activation function is softmax [9].

VI. VGG16

VGG16 is also a convolutional neural network. It is an image classification neural network that uses 3*3 filters one after the another. As compared to inception the hidden layers are less than the inception network. Also, a batch normalization layer is added to the VGG16 network[10] not only for overfitting but it provides the dataset batches one by one while the training period. The main concept of VGG16 is that it does not use high-level features making it a network for easy classification. It provides mainly all types of classification models. Fig. 6 shows the different layers in VGG16 mainly it comprises of convo 1-1 layers and convo 1-2 followed by the pooling layers. These convo layers keep changing in their instances. As there are dense layers in the end the final output is carried by the dense layer which is then taken for validation or the accuracy level.

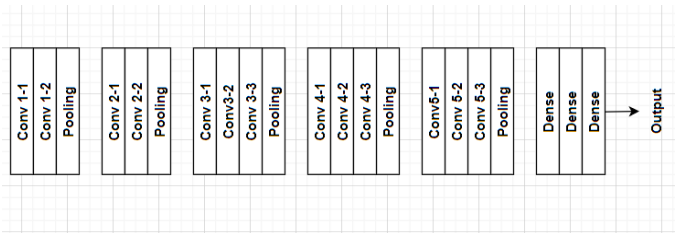

Fig. 6. VGG16 layers

In VGG16 two activation function [9] are used which are relu and softmax. The mathematical equation to calculate softmax function is:

$$\sigma(z^{\rightarrow})_i = \frac{e^{z_i}}{\Sigma_{j=1}^{K}\ e^{z_j}}$$

Where all symbols represent their own meaning[11],

σ = softmax

$z^{\rightarrow}$ = input vector

e^{z_i} = standard exponential function

K = number of classes in the multi class classifier

e^{z_j} = exponential function for output vector

VII. RESULTS

After the training of both models, it has been observed that both the models were trained very successfully and provides their individual training accuracy. Also, during the training, we have changed the batch size to 12 so that the maximum of images can be moved to the training models. Moreover, the epochs were also increased to a certain limit and it was found that the accuracy was also increasing and the production was at its best.

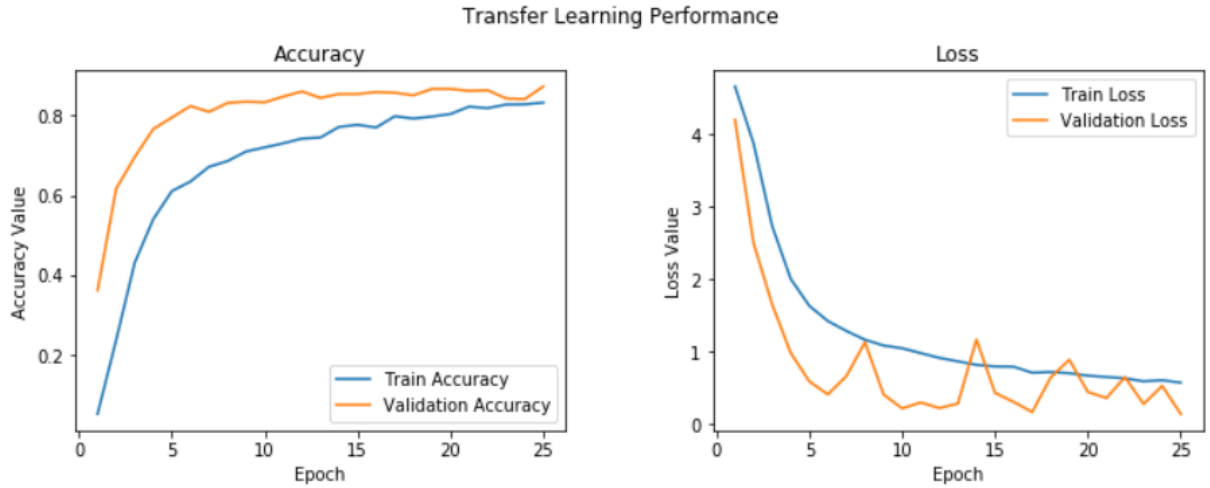

Fig. 7. Accuracy and train loss graph for Inception V3 neural network

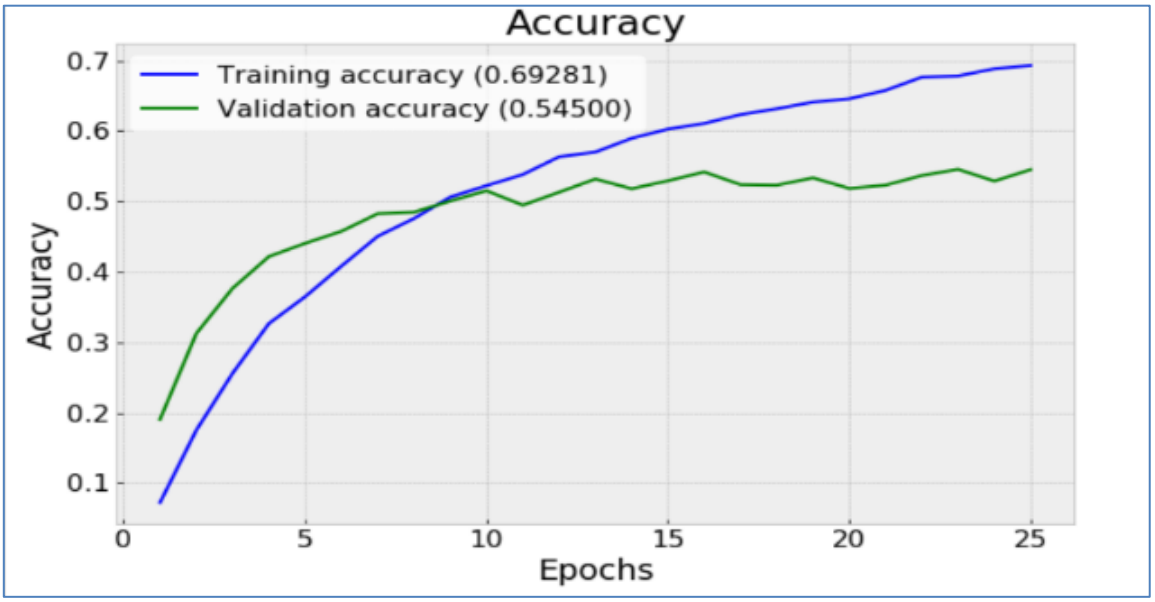

Fig. 8. Accuracy graph for VGG16 neural network

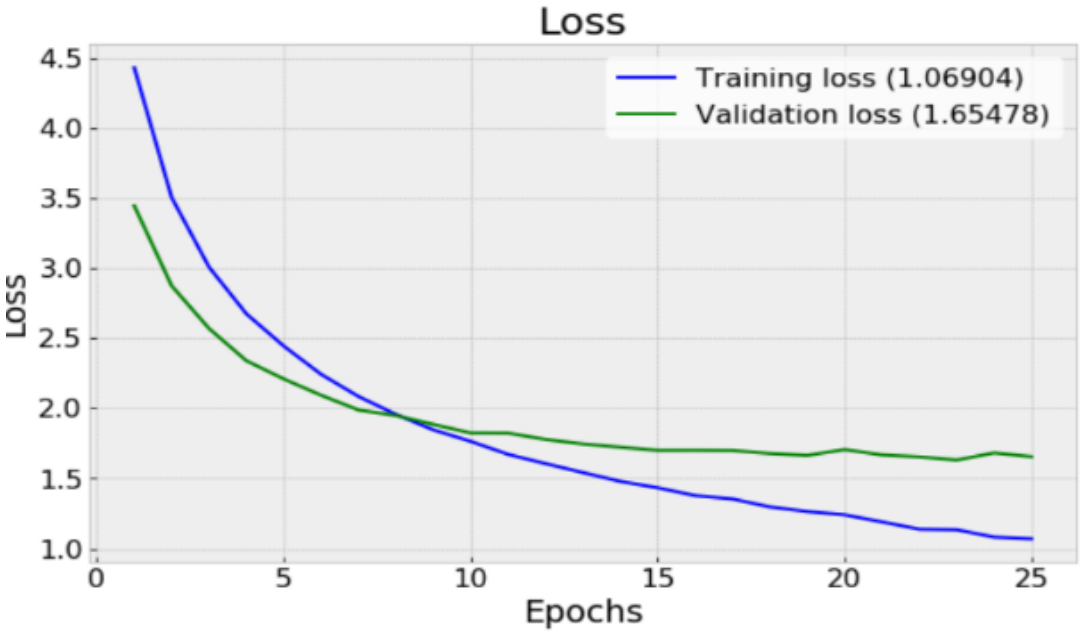

Fig. 9. Loss graph for VGG16 neural network

From Fig. 7 shows the transfer learning performance of the Inception V3 neural network. The first graph shows relationship between accuracy value and number of epochs which contains train accuracy and validation accuracy with the maximum accuracy of 85 with epochs equals to 25. Similarly, second part of this graph shows the relationship between loss value and number of epochs which shows train loss and the validation loss in the model. After evaluating the graph, it was observed that as the epochs increases the accuracy also increases and the loss value decreases.

Similarly, Fig. 8 and 9 shows graph for VGG16 which provides relationship between accuracy v/s epochs and loss v/s epochs accuracy observed in this model is 69 with a loss of 1.06.

VII. CONCLUSION

The paper provides comparison between Inception V3 and VGG16. Since there was also the comparison made between the two models it was done based on all the same parameters i.e., same number of epoch and same number of group size and the training data set was also same. It has been observed that Inception V3 provides an accuracy of 85 and on the other hand VGG16 provides an accuracy of 69 which is far more less than the Inception V3 model. Since the epoch provided to both the model were same and the classes were also same. new breeds are produced which is not able to identify. Overall a productive comparison has been made between both the models.

After the completion of this study, we came across many of the positive results;

1. As compared between both of the models Inception V3 has much higher accuracy then the VGG16.
2. Train loss as compared is also less in Inception V3 then VGG16.

REFERENCE

[1] P. Borwarnginn, K. Thongkanchorn, S. Kanchanapreechakorn, and W. Kusakunniran, "Breakthrough Conventional Based Approach for Dog Breed Classification Using CNN with Transfer Learning," *2019 11th Int. Conf. Inf. Technol. Electr. Eng. ICITEE 2019*, vol. 7, pp. 1–5, 2019, doi: 10.1109/ICITEED.2019.8929955.

[2] J. Liu, A. Kanazawa, D. Jacobs, and P. Belhumeur, "Dog breed classification using part localization," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7572 LNCS, no. PART 1, pp. 172–185, 2012, doi: 10.1007/978-3-642-33718-5_13.

[3] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar, "Localizing parts of faces using a consensus of exemplars," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 545–552, 2011, doi: 10.1109/CVPR.2011.5995602.

[4] M. Chanvichitkul, P. Kumhom, and K. Chamnongthai, "Face recognition based dog breed classification using coarse-to-fine concept and PCA," *2007 Asia-Pacific Conf. Commun. APCC*, pp. 25–29, 2007, doi: 10.1109/APCC.2007.4433495.

[5] D. Hsu, Stanford University "Using Convolutional Neural Networks to Classify Dog Breeds," pp. 1–6, 2015.

[6] Jesse Candido Standford University,"Dog Breed Classification and Visualization", CS230, Deep Learning.

[7] "Xiaolong Wang , Vincent Ly , Scott Sorensen and Chandra Kambhamettu Video / Image Modeling and Synthesis Laboratory , Department of Computer and Information Sciences , University of Delaware , Newark , DE," *Int. Conf. Image Process.*, pp. 5237–5241, 2014.

[8] B. A. M. Ashqar and S. S. Abu-Naser, "Identifying Images of Invasive Hydrangea Using Pre-Trained Deep Convolutional Neural Networks," *Int. J. Control Autom.*, vol. 12, no. 4, pp. 15–28, 2019, doi: 10.33832/ijca.2019.12.4.02.

[9] J. Ngiam, D. Peng, V. Vasudevan, S. Kornblith, Q. V. Le, and R. Pang, "Domain adaptive transfer learning with specialist models," *arXiv*, 2018.

[10] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection," *Constr. Build. Mater.*, vol. 157, no. September, pp. 322–330, 2017, doi: 10.1016/j.conbuildmat.2017.09.110.

[11] M. Wang, S. Lu, D. Zhu, J. Lin, and Z. Wang, "A High-Speed and Low-Complexity Architecture for Softmax Function in Deep Learning," *2018 IEEE Asia Pacific Conf. Circuits Syst. APCCAS 2018*, pp. 223–226, 2019, doi: 10.1109/APCCAS.2018.8605654.