Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 9, Issue. 4, April 2020, pg.76 – 82

Dog Breed Identification Using Pre-Trained Models

Tejeswar Sadanandan[a]; Thaha Mohammad S.Y[b]; Subhash Chandar T.J[c]; Vaidhya G.K[d]
[a]UG Student, CSE & Rajalakshmi Engineering College, Tamil Nadu, India
[b]UG Student, CSE & Rajalakshmi Engineering College, Tamil Nadu, India
[c]UG Student, CSE & Rajalakshmi Engineering College, Tamil Nadu, India
[d]Asst. Prof, Dept. of CSE &Rajalakshmi Engineering College, Tamil Nadu, India
tejeswar.s.2016.cse@rajalakshmi.edu.in; thahamohammed.sy.2016.cse@rajalakshmi.edu.in;
subhashchandar.tj.2016.cse@rajalakshmi.edu.in; vaidhyagk@gmail.com

Abstract-- Dog breed identification problem has many inter-class similarities and makes it difficult to classify or identify dogs which have similarities, we have about 130 different dog breeds and they are identified by convolutional neural network using transfer learning pre-trained models ResNet50 to obtain an accuracy of 82.7%, we use OpenCV implementation to detect human faces and to predict the nearest resemblance to a breed of dog, so if misclassified as human the predicted breed can be known. We use 30 epochs and a batch size of 64 to achieve this accuracy.
Keywords-- Dog breed identification, convolutional neural network, ResNet50

I. INTRODUCTION

Dogs are considered as man's best friend and lifelong companion and they are widely considered as the best animal to pet and can be easily trained and used as service dogs to aid the handicapped and guide them. They are used in police forces. There are over 340 dog breeds known throughout the world [1]. Dogs require atmost care and attention and their food and habitat environment depends on their breed and hence we have to identify their breeds to take care of them.

Expert advice and DNA tests can be taken to know the dog breed , but it is difficult to find an expert and it is a painful process for a dog and an expensive one too. And so we use a dog breed identification system to identify dog breeds and to provide necessary care. It is difficult even for experts to recognize a dog by its breed as many dog breeds recognized by AKC have similar facial features and it makes it hard for an expert to predict the breed of a given dog, ex: Curly Coated Retriever, American Water Spaniel have similar facial features such as ears, fur, nose but they are two totally different breeds.

In this project we identify dogs using machine learning such as Convolutional Neural Networks (CNN) and it is done by using transfer learning [2]. We used a pre-trained ResNet50 model to detect dogs in images. The ResNet50 model, along with weights that have been trained on ImageNet, a very large, very popular dataset used for image classification and other vision tasks. ImageNet contains over 10 million URLs, each linking to an image containing an object from one of 1000 categories. Unlike traditional sequential network architectures such as AlexNet, OverFeat, and VGG, ResNet is instead a form of "exotic architecture" that relies on micro-architecture modules (also called "network-in-network architectures").

The term micro-architecture refers to the set of "building blocks" used to construct the network. A collection of micro-architecture building blocks (along with your standard CONV, POOL, etc. layers) leads to the macro-architecture (i.e,. the end network itself)[3].

Figure 1. High inter-class correlation

II. RELATED WORK

This section of the paper attempts to capture previously faced problems by tackling current research. Kumar et al. [4] implemented a fusion based method with 94.86% rank-4 identification rate and later they improved their accuracy to 96.87% via a Fisher Linear Projection and Preservation approach [5] but the dataset they used is hidden and so the high accuracy obtained cannot be compared with others. Liu et al. [6] created a dataset of images of various breeds. To classify dog breeds, they detected dogs with a SVM regressor and located the region of interest using an SVM based sliding window detector, their accuracy was 67.00%.

Zhang et al. [7] used RCNN to detect parts of an image. Angelova et al. [8] used object detection and segmentation for this research. that describes an approach to identify dogs using four existing human facial recognition methods (EigenFaces, FisherFaces, LBPH, and a Sparse method). Two CNNs were also tested on the dataset of two selected breeds: huskies and pugs, with accuracy of 75.14% for huskies and 54.38% for pugs.

Howard et al. [9] proposed a model based on depth-wise separable convolutions with 83.30% accuracy. An attention-based model was applied in [10] with 76.80% accuracy. The DenseNet model for breed classification was applied in [11] with 85.14%accuracy.

III. DATASET AND PREPROCESSING

The method in this paper is CNN using transfer learning where the dataset from [12], it consists of dog images of 133 different dog breeds. It is separated into training data and test data and validation data. The dataset consists of a total of 8,351 images with 6,700 training images with 825 validation images and 826 test images and 13,233 total human images as we identify dogs only if they are identified to be dogs if not we see if the given image is resembling human and nearest predicted dog breed.

The preprocessing is done by splitting the dataset into train, test and valid. The training dataset contains roughly 50 dog images for a particular breed. The test dataset contains roughly 8 dog images for a particular dog breed. The validation dataset contains roughly 6 images per breed.

The images are converted to an array and resized as 4D tensor but first they are reshaped into a square image of 224x224 pixels. Each image has 3 channels for colour images.

F**igure 2.** Sample images in dataset

IV. METHODOLOGY AND EXPERIMENTS

A. Predictions with ResNet50:

Some additional processing is required for getting the 4D tensor ready for pre-trained models like ResNet50. The image is converted from RGB to BGR image by reordering the channels. Every mean pixel should be subtracted from every pixel of an image, this is a required additional step for normalization for all pre-trained models.

The model is ready to be used to extract predictions, since we have a way of formatting the image to be supplied to ResNet50. An array whose ith entry is the model's probability predicted that image belongs to ith ImageNet category and it is returned from the predict method which does the required task. Relevant Integer is obtained to the model's predicted class by using argmax of the predicted vector probability.

B. Test human face detector:

Ideally, we would like 100% of human images with a detected face and 0% of dog images with a detected face. You will see that our algorithm falls short of this goal, but still gives acceptable performance. Numpy arrays are used to store the file paths after we extract file paths of 100 first images from our dataset. After testing you can see that 99% of the first 100 images of humans have a detected human face, while 12% of the first 100 images of dogs have a detected human face.

C. Test the dog detector:

As expected from our detector, 0% of the images of humans detected a dog and 100% of the images of dogs have a detected dog. This proves that the dog detector we constructed is working well. To calculate the features all possible sizes and each kernel location are used.(Even a 24x24 window requires a lot of computation and produces results over 160000 features). We need to find the sum of pixels under white and black rectangles for each feature calculation. They introduced integral images to solve this. How large may be the number of pixels, to an operation involving just four pixels and it simplifies calculation of sum of pixels A window of the target size is moved over the input image during the detection phase and haar features are calculated for each subsection of the image. A learned threshold that separates non-objects from objects is used to compare the difference. Haar features are only a weak classifier (detection quality is only a bit better than random guessing) so large numbers of features are required to describe objects with good accuracy and they are organized to cascade classifiers to form strong classifiers.

D. Haar cascade:

Haar Cascade is used to identify objects present in an image or motion video and it is an object detection algorithm proposed by by Paul Viola and Michael Jones in their paper "Rapid Object Detection using a Boosted Cascade of Simple Features" in 2001.[13]

From a lot of positive and negative images, Haar cascade which is a machine learning approach where the function is trained then the objects are detected from images.

The algorithm stages are:

Haar Feature Selection

Creating Integral Images

Adaboost Training

Cascading Classifiers

It can be used to train and identify any object and it is also widely used for its capability to detect human faces and other body parts in any given image. Consider the image given is an example of a human face and the algorithm starts by its requirement for many positive images including human faces and negative images also should be provided which has no human faces for the classifier to be trained. Further the features are extracted from it.

At the beginning of the process Haar Features are collected. Adjacent rectangular regions at a specific area are considered in the detection window, the difference is calculated later by adding the pixel intensities in our regions.

To calculate the features all possible sizes and each kernel location are used.(Even a 24x24 window requires a lot of computation and produces results over 160000 features). We need to find the sum of pixels under white and black rectangles for each feature calculation. They introduced integral images to solve this. How large may be the number of pixels, to an operation involving just four pixels and it simplifies calculation of sum of pixels A window of the target size is moved over the input image during the detection phase and haar features are calculated for each subsection of the image. A learned threshold that separates non-objects from objects is used to compare the difference. Haar features are only a weak classifier (detection quality is only a bit better than random guessing) so large numbers of features are required to describe objects with good accuracy and they are organized to cascade classifiers to form strong classifiers.

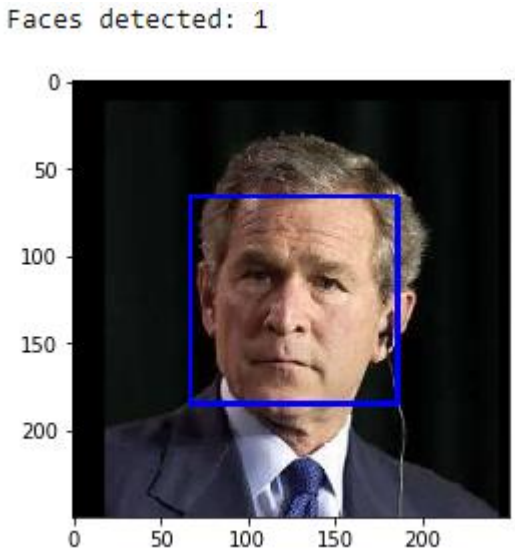


Figure 3. Face detection using Haar Cascade

V. FINAL RESULT

This experiment focuses on finding or identifying the dog breed and to differentiate it from human faces. The experiment uses ResNet50 and uses predefined models and weights to predict the dog breeds which are input. The experiment uses a human face detector and dog detector and the results are as follows

Percentage of humans correctly classified as people: 99.0%
Percentage of dogs misclassified as people: 12.0%
Percentage of humans misclassified as dogs:: 1.0%
Percentage of dogs correctly classified as dogs: 100.0%

The algorithm has an accuracy of 82.7% by using CNNs (using transfer learning). At first the model trains on 30 epochs with a batch size of 64 and then it is done with 5 epochs to obtain this accuracy

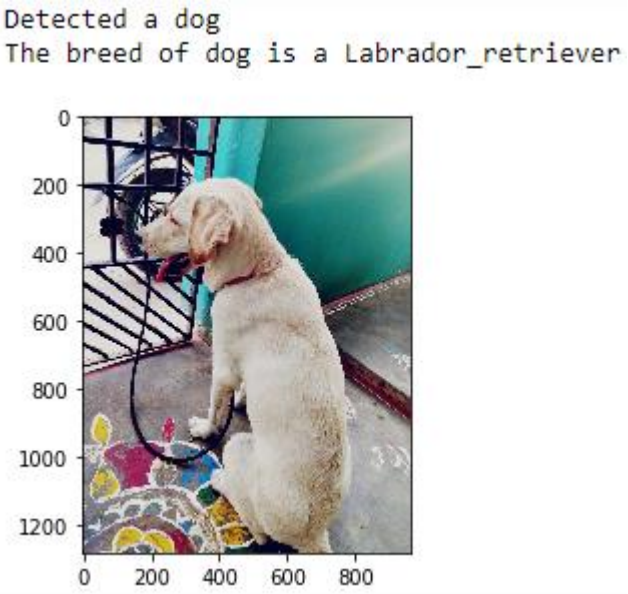
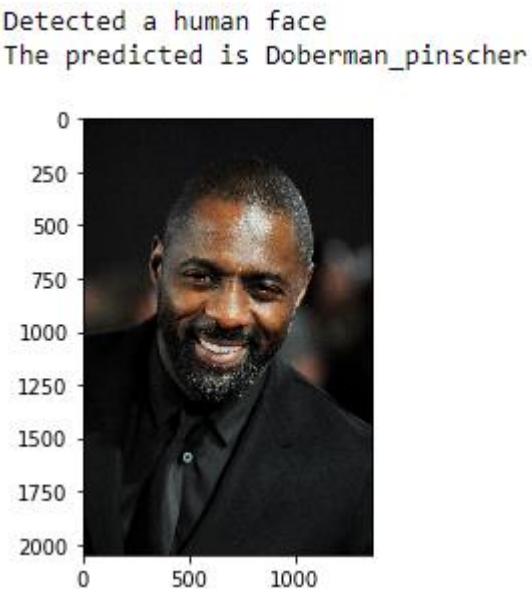


Figure 4. Dog image is classified and breed is identified correctly



F**igure 5.** Human image classified correctly as human and the predicted breed is shown

VI. CONCLUSION

This paper presents identification of dogs and differentiates them from humans using pre trained CNNs. There are 133 breeds of dogs and they are predicted with an accuracy of 82.7% using the ResNet50 algorithm using CNN (transfer learning) which is provided by keras and the accuracy is found to be highest among all the other algorithms which use keras. The system specification of the computer used to run the dog breed identification program is as follows

Name	Specification	Model
CPU	Intel	i5 5200U
GPU	Intel HD graphics	5500
	DDR3	8GB
OS	Windows	10

Table 1. System Specification

REFERENCES

[1] AKC American Kennel Club, https://www.akc.org/dog-breeds/.
[2] J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.
[3] Pyimagesearch, https://www.pyimagesearch.com/2017/03/20/imagenet-vggnet-resnet-inception-xception-keras
[4] S.Kumar and S.K.Singh. Biometric recognition for pet animal. Journal of Software Engineering and Applications,7(05):470, 2014.
[5] S.Kumar and S.K. Singh.Monitoring of pet animals in smart cities using animal biometrics. Future Generation Computer Systems, 83:553-563, 2016
[6] J. Liu, A. Kanazawa, D. Jacobs, and P. Belhumeur. Dog breed classification using part localization. In European Conference on Computer Vision, pages172-185.Springer, 2012.
[7] N. Zhang, J. Donahue, R. Girshick, and T. Darrell, "Part-based r-cnns for fine-grained category detection," in Computer Vision – ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 834–849.
[8] *A.Angelova and S.Zhu, "Efficient object detection and segmentation for fine-grained recognition," in Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, 2013, pp. 811–818.*
[9] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861,2017.
[10] P. Sermanet, A. Frame, and E. Real. Attention for fine-grained categorization.arXiv preprintarXiv:1412.7054,2014.

[11] W.Shi,J.Chen,F.Liu,and F.Liu. Dog breed identification.
[12] Udacity Data Science nano degree, https://www.udacity.com/course/data-scientist-nanodegree--nd025
[13] Rapid Object Detection using a Boosted Cascade of Simple Features" 2001 Paul Viola and Michael Jones.

