International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 06 | June - 2021 ISSN: 2582-3930

IMAGE CLASSIFICATION (CATS AND DOGS) WITH DIFFERENT CONVOLUTIONAL NEURAL NETWORKS

A1: Abhilash R

M.C.A. Student, Dayananda Sagar College of Engineering

A2: Amith B N

Assistant Professor, Department of MCA, Dayananda Sagar College of Engineering

ABSTRACT

Image classification has become more popular as it is the most basic application and implementation of deep learning. Images of dogs and cats are the most common example to train image classifiers as they are relatable. In this paper, we tried to build an image classifier to recognize various pictures of dogs and cats (CDC) using one own neural network and one fine-tuned VGG16 to build the classifier. The resulting model from own neural network has a training accuracy of 98.47%, validation accuracy of 60%. The model from VGG16 has a training accuracy of 99.59%, validation accuracy of 96.56%.

In this model it has several steps ,first we train the model on different patterns like using a own neural network model and pre-trained VGG16 model , second we predict the given new image to check with cat or dog ,at last we also use gradient cam technology to check how the model comes to the conclusion of decision that is on what basis it outputs results.

1.INTRODUCTION

The Dogs vs. Cats image classification has been around for a long time now. The Dogs vs. Cats competition from Kaggle is trying to solve the CAPTCHA challenge, which relies on the problem of distinguishing images of dogs and cats. It is easy for humans, but evidence suggests that cats and dogs are particularly difficult to tell apart automatically.

The convolution neural network (CNN) developed in recent years has been widely used in the field of image processing because it is good at dealing with image classification and recognition problems and has brought great improvement in the accuracy of many machine learning tasks. It has become a powerful and universal deep learning model.

Many people has worked or are working on constructing machine learning classifiers to address this problem. A classifier based on color features got 56.9% accuracy on the Asirra dataset. An accuracy of 82.7% was achieved from a SVM classifier based on a combination of color and texture features.

In my project I am have built a convolutional neural

network to solve the problem and achieve higher performance and better results . In my project instead of using complete Kaggle data set comprising of total 25000 images, I have worked on subset of these images. My dataset would be comprising of total 10000 images.

Keras would used for model building and all the code would be implemented on jupyter notebook.

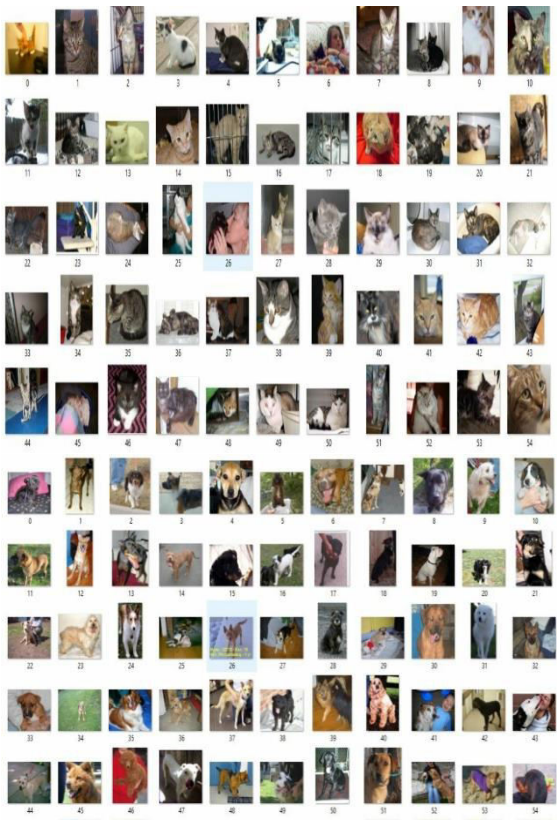
As the trial of comparing a simple neural network with fine tuned and huge VGG16 neural network we took vgg16 for learning purpose as it is easy to implement among other neural network . I this we took Kaggle dataset images for training and testing and will try to compare both the networks by performing similar steps

depending on the model compatibility. Further in this paper we will discuss all the testing and comparison criteria with the associated results.

2.DATASET AND EVOLUTION

The Kaggle dataset contains 25,000 images we can download the dataset using this link

https://www.microsoft.com/enus/download/details.aspx?id=54765



- In order to handle data in own neural network I have used pickle to load and save dataset.
- In VGG16 model keras data generators handled the dataset.
- Open CV has worked in managing image operations.

GradCAM function for producing heat maps visualisations.

3.ABOUT MODELS

3.1. Own neural network

This convolutional neural network contains 2 layers of conv2D and 2 MaxPooling with the activation function relu ,a flatten layer followed by this 2 dense layers one with activation function relu and other with softmax.

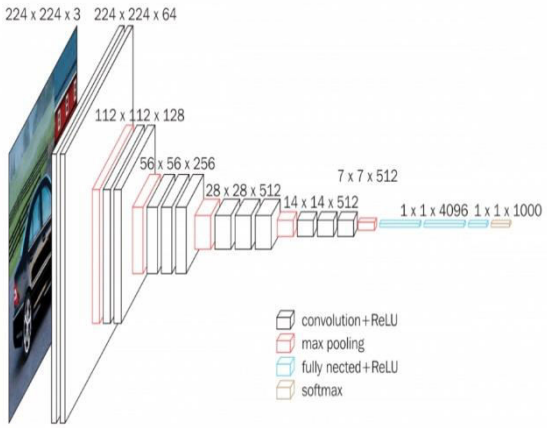
AnAdam optimizer and sparse_categorical_entropy is used as loss function.

3.2. VGG16 neural networks

VGG16 is used in many deep learning image classification problems; however, smaller network architectures are often more desirable. But it is a great building block for learning purpose as it is easy to implement.

This network consist of 3*3 convolutional layers one after the other with activation functions RELU and one outmost layer with softmax as per the below figure.

We can use Imagenet (large visual database designed for use in visual object recognition) which are pre-trained weights for VGG16 which helps in prediction of images.



3.3.Convolutional Neural Network

Convolutional Neural Networks have a different architecture than regular Neural Networks. Regular Neural Networks transform an input by putting it through a series of hidden layers. Every layer is made up of a set of neurons, where each layer is fully connected to all neurons in

the layer before. Finally, there is a last fully-connected layer — the output layer — that represent the predictions.

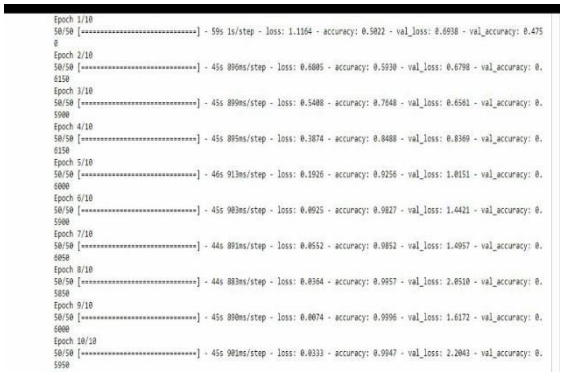
4.RESULTS AND ANALYSIS

4.1Training results of created neural network

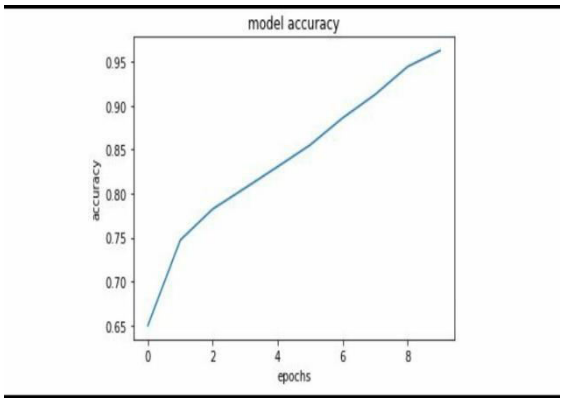
Model.fit function results

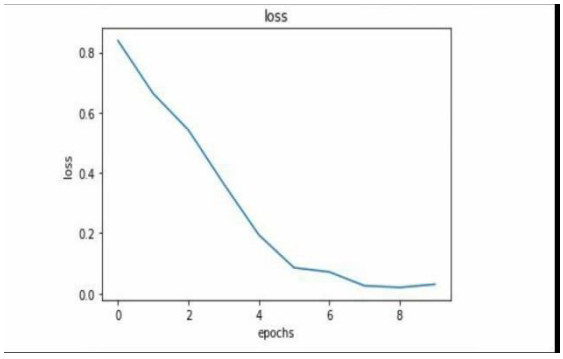
We have used 10 epochs to run the training sessions with 50 steps per each epoch,

Every epoch returns the value of accuracy,loss , validation accuracy and validation loss,This helps us to know the metrics of training.



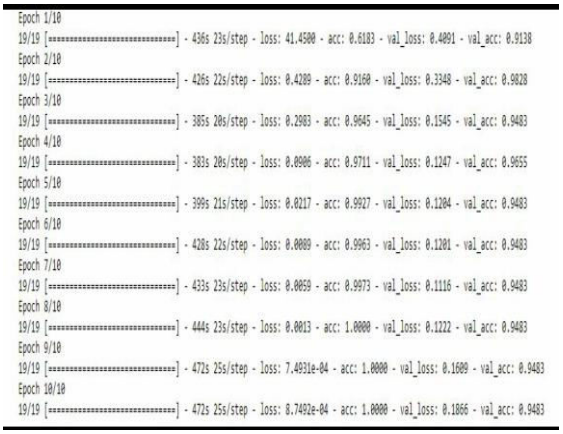
Graphs of the above results

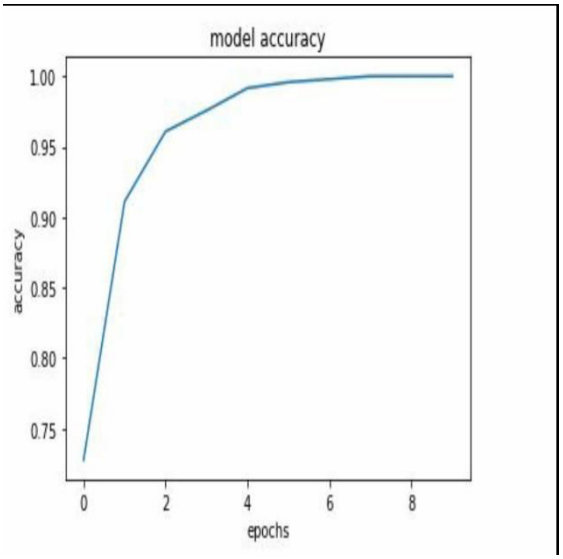


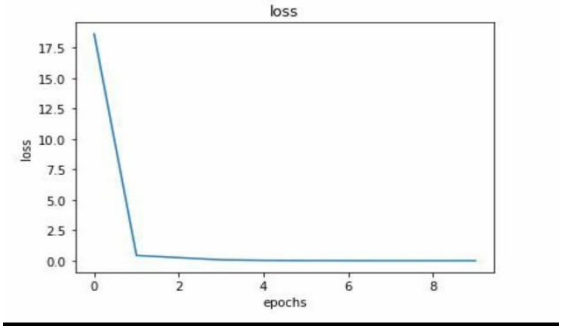


4.2 Training results of VGG16

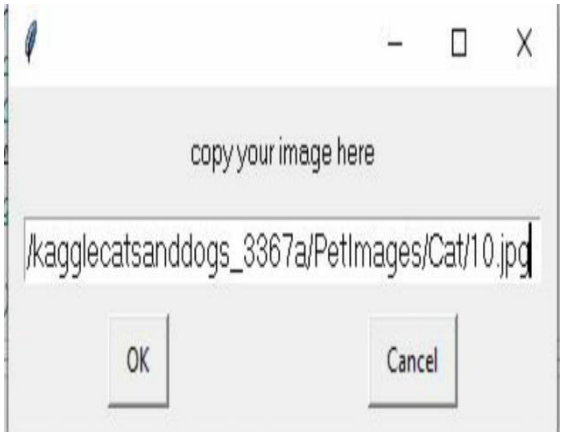
We will be performing same training process with metrics accuracy,loss, validation accuracy and validation loss. VGG16 attains 100% accuracy at the early epochs and maintains the same.



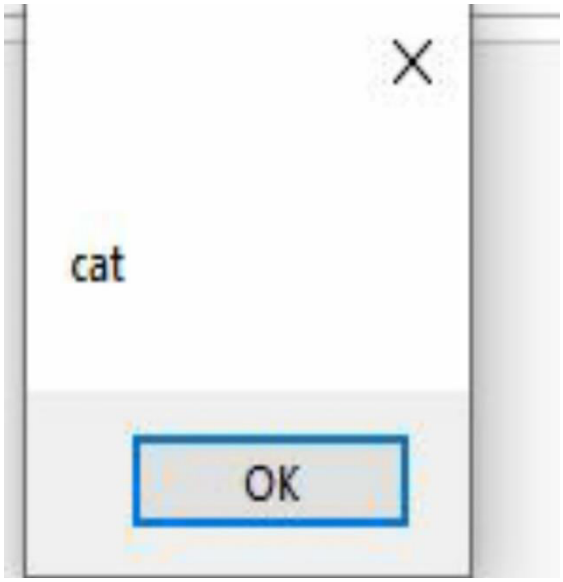




4.3Inference function (prediction function)

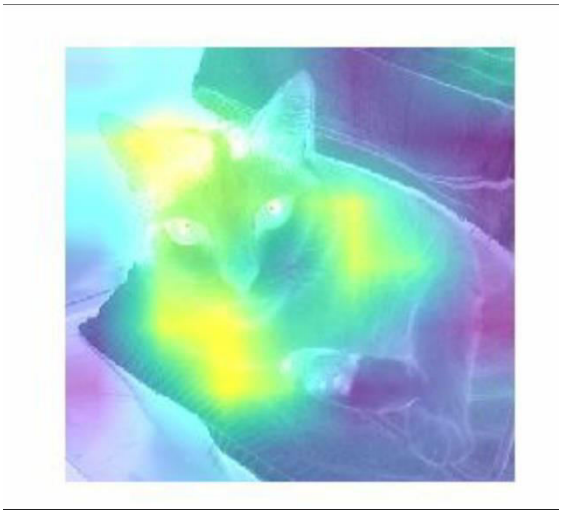


OUTPUT

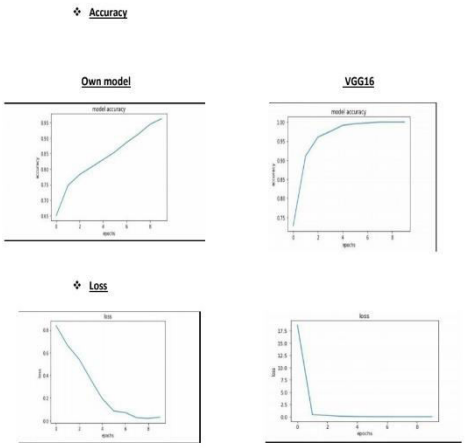


4.4 Heatmap (GradCAM)

Grad cam function has used to create heat maps in order to check how neural network analyzes the given picture and process the results, the greenish regions are the key features considered for analysing the picture as cat .



5.Comparing created nueral network and vgg16



6. Comparison table

Category	Own network	VGG16
Accuracy	99.45%	100%
Loss	3%	.87%
Validation accuracy	60%	94.8%
Time taken for 5 epochs (training)	835secs	1582secs
Inference time	.55secs	1.1secs

7. Conclusion

- We are able to achieve 93 percent accuracy when taken average on both the network models .
- The created neural networks fits more efficiently than vgg16 as vgg16 overfitsl.
- Vgg16 model slightly overfits as graphs have much variations.
- VGG16 is heavy Neural network and consumes more cpu functioning and needs a high Ram availability and it takes more time when compared to the other own neural network, even though own network records some less accuracy than vgg16 we can get the same results using own network.
- Own network works more efficiently with lower system requirements and it faster than vgg16 so that we can use own neural at the place of fine tuned VGG16 ,when it comes systems at lower efficiency or availabilities.
- Models works efficiently in creating heat maps .
- This was a trial to understand convolutional neural networks and have successfully understood the performance differences and need of neural networks.

8. References

• Microsoft.com for downloading the Dataset

• Kaggle

• Tensorflow.org

•Sentdex – YouTube channel for understanding concept

• Stackoverflow.com