Skin Cancer Classification using VGG-16 and Googlenet CNN Models

Suryakanth B. Ummapure, PhD
Department of Computer Science
Govt. First Grade College
Shahapur, Karnataka, India

Ravindrakumar Tilekar
Department of Computer Science
Govt. First Grade College Bidar, Karnataka, India

Satishkumar Mallappa, PhD
Sri Satya Sai University for Human Excellence Kalaburagi Campus, Navanihal , Kamalapur Karnataka

ABSTRACT
Skin cancer with a high fatality rate is called melanoma. Due to the great degree of similarities among the many forms of skin lesions, a proper diagnosis cannot be made. Dermatologists can treat patients and save their lives by accurately classifying skin lesions in their early stages. This paper proposes a model for highly accurate skin lesion classification. The proposed model made use of transfer learning models known as GoogleNet and vgg16. This model efficiently distinguished between benign and malignant cancerous skin lesions, those are the two distinct classes of skin diseases. The 1800 benign cancer images and 1498 malignant cancer images that were retrieved from the internet were taken into account for this proposed strategy. The VGG16 has obtained the highest recognition accuracy in the result accessing, with recognition rates of 99.62% for training and 84.97% for validation.

Keywords
Googlenet, VGG16, Skin Cancer, Deep Learning

1. INTRODUCTION
Skin cancer [1] [2] is a form of cancer that arises when DNA damage that has not been repaired causes abnormal growth in skin cells. It is best to identify skin cancer early since it is more manageable in its initial stages, despite the fact that it spreads progressively to other regions of the body. Early detection of skin cancer signs is necessary due to the increased rate of the disease, high death rate, and high expense of medical care. The two types of skin cancer are classified as malignant and benign. Traditional computer vision methods are frequently used as a classifier to extract numerous features, such as shape, size, color, and texture, in order to identify cancer. Due to the availability of deep learning techniques, these features are now not very frequently utilized. Feature classification and feature detection are two most common deep learning methods. This method is capable of extraction the potential features of its own and those features are called CNN features. Hence, from skin cancer images, CNN features were extracted and efficiently utilized for the classification purpose.

The flow of this paper is in this manner section-2 describes earlier works carried out in the relevant area. The experimental details are given in section-3. The outcome of the experimental results and discussions are presented in section-4. At last, section 5 ends the paper with conclusion and future scope of this work.

2. PREVIOUS WORK
If skin cancer is found in its early stages, treatment may be given. The patient's skin observations can aid surgeons in making the best treatment options based on skin cancer images. Numerous studies on the identification and classification of skin cancer have been published in the literature. The paper published by Shahin et.al[3], have proposed the deep learning based skin cancer classification they have used HAM10000 dataset and obtained the 96.16% recognition accuracy for training and for testing they have achieved the 91.96% recognition accuracy respectively and they have compared the model with other models such as Resnet, AlexNet, VGG-16,MobileNet DenseNet, etc.

In Jinnai et.al.[4], they have reported the work on skin cancer development for the pigmented skin lesions by applying the deep learning method from this they have got an highest of 91.5% highest recognition accuracy. The integrated design of deep features fusion based skin cancer classification is reported by Amin et al.[5], they have used the deep features and applied PCA on the features and obtained the 99.00% recognition accuracy.

An seven ways skin cancer classification by using Mobilenet is reported by Chaturvedi et al.[6], by using pretrained model named as Mobilenet on 2014 ImageNet Challenge dataset and obtained the overall accuracy of 83.10%. A review paper on skin cancer classification by using various methods; deep learning, and CNN is given by Manne et al.[7].

The interpreted deep learning method to segment and classification of non-melanoma skin cancer is reported by Thomas et al.[8], from this method they have obtained the 97.9% recognition accuracy. Filali et al.[9] is shown the best use of hand crafted and CNN features for skin cancer classification and they have used the PH2 dataset and obtained the 98% recognition accuracy.

A novel approach is reported in the work of Singh et al.[10], they have reported the transfer learning (TL) framework Transfer Constituent Support Vector Machine (TrCSVM) from this method they have obtained the 98.82% overall recognition accuracy.

3. SIMULATION DETAILS
In the process of carrying out this proposed experiment the two popular pretrained CNN models; Googlenet and VGG16 applied on 3298 skin cancer images. The pretrained models are widely used to carry out the image classification kinds of work. To obtain the maximum recognition accuracy these models may be used. These are the models which are built by someone else and others were used to test their data on them with the intention of getting maximum recognition accuracy to fulfill the research objective belongs to many research areas. By using these pretrained models for others experiment is sometime known as transfer learning also.

In these pretrained models the layers were efficiently utilized, in the matter of numbers and input parameters were considered well. The main objectives about these models are the size and training time taken by the pretrained model. This proposed experiment have been utilized the widely used pretrained CNN models namely, Googlenet and VGG16

Googlenet: This network is first reported by Szegedy et al.[11] , mainly this network is used for classification and detection the objects from the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14). It employs a variety of techniques, including global average pooling and 1-

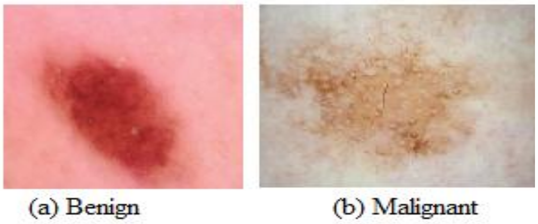

Fig 1: Sample Skin cancer images

1 convolution, to build deeper architecture. The 1-1 convolution is a feature of the inception architecture. The number of parameters (weights and biases) in the architecture was reduced using these convolutions. The majority of parameters in many architecture that increase computation cost are found in these fully connected layers. It is an 22 layered network. Following figure shows the detailed architecture of the Googlent.

The organization was planned considering computational effectiveness and practical sense, so this network can be run on single CPU incorporating even those with restricted computational assets, particularly with low-memory machines. In this proposed method we have made use of this network to classify the skin cancer images. This model has successfully achieved the more accuracy by efficiently achieving the input image size reduced and retained the important features in the image. This network has embedded with additional component named as auxiliary classifier. This classifier is used in the training phase only and it is eliminated in the inference. The best part of this auxiliary classifier is that it performs the classification process on input images in the network midsection and it adds the well calculated loss in the training process.

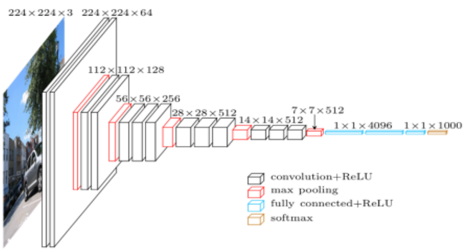

Fig 2: VGG16 architecture

VGG16: Initially this network is proposed by Simonyan et al. [12], it is one of the best network models till date. Beginning with VGG11 and ending with VGG19, the VGG research group released a series of convolution network models. The 1800 benign cancer images and 1498 malignant cancer images that were retrieved from the internet were taken into account for this proposed strategy[13].

The primary goal of the VGG group on depth was to comprehend how the accuracy of the models for large-scale image classification and recognition is affected by the depth of convolutional networks.

In comparison to the maximum VGG19, which has 16 convolutional layers and the three fully connected layers, the minimum VGG11 has 8 convolutional layers and three fully connected layers. The final three fully connected layers have exactly the same variations of VGGs.

Five sets of convolutional layers make up the overall structure, which is followed by a MaxPool. The difference is that in the five sets of convolutional layers, as the depth rises—that is, as it go from VGG11 to VGG19—more and more cascaded convolutional layers are added. The VGG16 refers to the 16 layers architecture. Following figure shows VGG16 architecture.

It is considered to be one of the excellent vision model architecture till date. The VGG16 has used the increasing depth by employing the small (3x3) convolution filters, this gives the network an significant improvements. This network can be used for large datasets.

In the experiment setup we have considered the 224x224 image sized skin cancer images as a input to the both the models. The dataset is divided in 80% and 20% margin. 80% dataset is used for training purpose and 20% dataset is used for testing purpose. The hyper parameters are set in this way, max epochs have set to 25, 128 mini batch size is fixed and 0.0001 learning rate with adam optimizer have been used. Analysis of the size reduction after each maximum pooling is essential in this case.

In ImageNet, a dataset that contains more than 14 million training images across 1000 object classes, the VGG16 model can achieve a test accuracy of 92.7%. It is a standout model from the 2014 ILSVRC competition. VGG16 enhances AlexNet by substituting sequences of smaller 33 filters for the large filters. For the first convolutional layer in AlexNet, the kernel size is 11, and for the second layer, it is 5.

Using NVIDIA Titan Black GPUs, the researchers trained the VGG model for several weeks. This model was different from earlier, successful models in a number of ways. The smallest 3x3 receptive field with a 1-pixel stride was used initially.

The function of a larger receptive field is achieved by the combination of the smaller 3x3 filters. When using multiple smaller layers as instead of a single large layer, the decision functions are improved and the network can converge more quickly. This is because there are more non-linear activation layers present.

The smaller convolutional filter used by VGG greatly reduces the possibility that the network will over-fit during training process. The best size for a filter is 3 x 3, as smaller sizes can't capture information from the left, right, and up and down. These models have suffering from two drawbacks those are this model slow to train and weights of this model is larger themselves.

4. EXPERIMENTAL RESULTS AND DISCUSSION

This section describes the outcome of our experiment used to classify the skin cancer types namely malignant and benign. The images are 1800 benign and 1498 malignant cancer images. Following figures shows the training plot, and confusion matrix for the Googlenet model.

Table 1. Confusion matrix for Training images

benign	1265	120
malignant	175	1078

Table 2. Confusion matrix for testing images

benign	296	66
malignant	63	233

Table 3. Confusion matrix for Training images from VGG16

benign	1436	6
malignant	4	1192

Table 4. Confusion matrix for Testing images from VGG16

benign	302	41
malignant	58	258

Table 5. Recognition accuracy of Googlenet and VGG16 models

Sl. No.	Models	Training Accuracy	Testing Accuracy
1	Googlenet	88.81%	80.54%
2	VGG16	99.62%	84.97%

Table 6. Googlenet archíture parameters

type	patch size/ stride	output size	depth	#1x1	#3x3 reduce	#3x3	#5x5 reduce	#5x5	pool proj	params	ops
convolution	7x7/2	112x112x64	1							2.7k	34M
max pool	3x3/2	56x56x64	0								
convolution	3x3/1	56x56x192	2		64	192				112k	360M
max pool	3x3/2	28x28x192	0								
inception(3a)		28x28x256	2	64	96	128	16	32	32	159k	128M
inception(3b)		28x28x480	2	128	128	192	32	96	64	380k	304M
max pool	3x3/2	14x14x480	0								
inception(4a)		14x14x512	2	192	96	208	16	48	64	364k	73M
inception(4b)		14x14x512	2	160	112	224	24	64	64	437k	88M
inception(4c)		14x14x512	2	128	128	256	24	64	64	463k	100M
inception(4d)		14x14x528	2	112	144	288	32	64	64	580k	119M
inception(4e)		14x14x832	2	256	160	320	32	128	128	840k	170M
max pool	3x3/2	7x7x832	0								
inception(5a)		7x7x832	2	256	160	320	32	128	128	1072k	54M
inception(5b)		7x7x1024	2	384	192	384	48	128	128	1388k	71M
avg pool	7x7/1	1x1x1024	0								
dropout (40%)		1x1x1024	0								
linear		1x1x1000	1							100k	1M
softmax		1x1x1000	0								

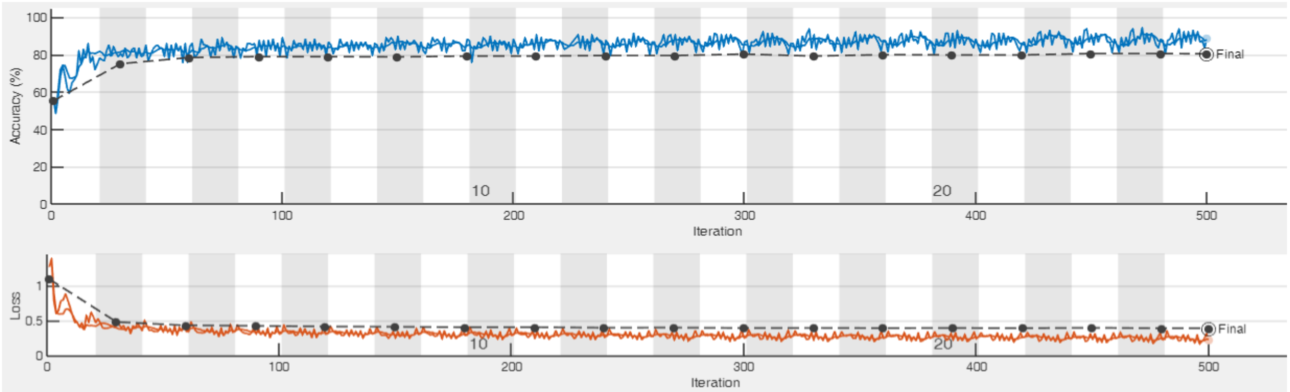

Fig 3: Training plot of Googlenet model

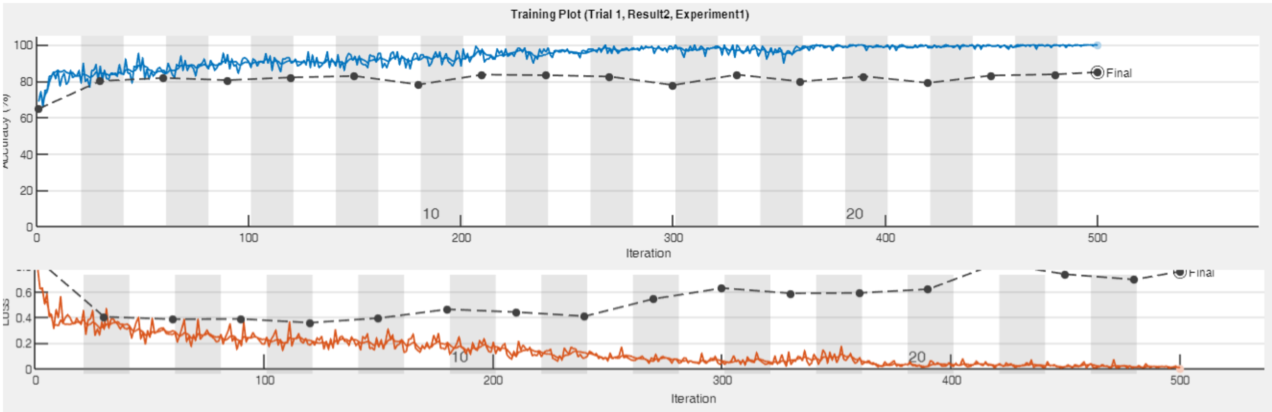

Fig 4: Training plot of VGG16

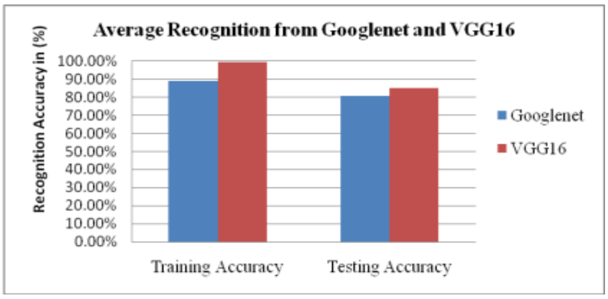

Fig 5: Skin cancer classification recognition accuracy

From the Figures and Tables it is shown the performance of Googlenet and VGG16 for classification of skin cancer types. The VGG16 has outperformed from the Googlenet by achieving highest 99.62% and 84.97% recognition accuracy obtained from the training and testing images.

5. CONSLUSION AND FUTURE SCOPE

In the rapid changes of the environment due to air pollution, water pollution and climate is unbalanced due to this many health issues are increasing rapidly. The unhealthy habits of human's leads to many diseases are starting. From these both polluted environment and unhealthy habits, many people's are suffering from lots of health issues.

In this proposed work the skin cancer is picked and studied and make effort to classify the cancer type from the images. The VGG16 model has reached to 99.62% this is near to 100% recognition accuracy.

In the future scope, the effort can be put to collect more sample images belonging to skin cancer by adopting other standard datasets available for free with this strategies are made to develop an own light weighted Deep learning model to classify the cancer images with highest accuracy.

6. REFERENCES

[1] Jinnai, S., Yamazaki, N., Hirano, Y., Sugawara, Y., Ohe, Y., & Hamamoto, R. (2020). The development of a skin cancer classification system for pigmented skin lesions using deep learning. Biomolecules, 10(8), 1123.

[2] Le, D. N., Le, H. X., Ngo, L. T., & Ngo, H. T. (2020). Transfer learning with class-weighted and focal loss function for automatic skin cancer classification. arXiv preprint arXiv:2009.05977.

[3] Md Shahin Ali, Md Sipon Miah, Jahurul Haque, Md Mahbubur Rahman, Md Khairul Islam," An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models", Machine Learning with Applications 5 (2021).

[4] Jinnai, Shunichi, et al. "The development of a skin cancer classification system for pigmented skin lesions using deep learning." Biomolecules 10.8 (2020): 1123.

[5] Amin, Javeria, et al. "Integrated design of deep features fusion for localization and classification of skin cancer." Pattern Recognition Letters 131 (2020): 63-70.

[6] Chaturvedi, Saket S., Kajol Gupta, and Prakash S. Prasad. "Skin lesion analyser: An efficient seven-way multi-class skin cancer classification using mobilenet." International Conference on Advanced Machine Learning Technologies and Applications. Springer, Singapore, 2020.

[7] Manne, Ravi, Snigdha Kantheti, and Sneha Kantheti. "Classification of Skin cancer using deep learning, ConvolutionalNeural Networks-Opportunities and vulnerabilities-A systematic Review." International Journal for Modern Trends in Science and Technology, ISSN (2020): 2455-3778.

[8] Thomas, Simon M., et al. "Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer." Medical Image Analysis 68 (2021): 101915.

[9] Filali, Youssef, et al. "Efficient fusion of handcrafted and pre-trained CNNs features to classify melanoma skin cancer." Multimedia Tools and Applications 79.41 (2020): 31219-31238.

[10] Singh, Lokesh, Rekh Ram Janghel, and Satya Prakash Sahu. "TrCSVM: a novel approach for the classification of melanoma skin cancer using transfer learning." Data Technologies and Applications (2020).

[11] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015

[12] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[13] https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign