An Efficient Framework using Deep Learning for Skin Cancer Classification

Sarthak Garg, Harshit Garg, Vikas Tripathi , Bhasker Pant, Kumud Pant

Abstract*: *A good image analysis model can be very helpful in accurate diagnosis/classification of diseases for which images are available. Due to a plethora of public image databases, training and testing of algorithms on the dataset have helped in the development of an efficient framework for image classification. Skin cancer is one such disease for which recently image databases have been developed. Of the various methods of classification of skin cancer based on image analysis, convolutional neural network (CNN) has proven to be better performing than conventional machine learning approach. Realizing the importance of developing an efficient framework for skin cancer classification, this paper proposes a framework which utilizes VGG-16 CNN model to classify cancer images into categories namely, malignant or benign. We have trained the model using the skin cancer images freely accessible on the ISIC archive and attained an accuracy of 97.81%.

Keywords: VGG-16 Model, Convolutional Neural Network, Deep Learning, ReLu, skin cancer, image processing.

I. INTRODUCTION

Being one of the most common type of cancer, skin cancer contributes to almost 40% of reported cancer cases worldwide. In malignant skin cancer, the cancer cells divide very rapidly and interfere with the normal functions of the cells of the body whereas in benign skin cancer the normal functioning of the cells of the body is not meddled with [1].

Benign tumors are normal cells that can divide at an alarming rate and grow quite large but they do not intervene with the functioning of the neighboring normal cells. They do not have the ability to migrate from where they originated. No matter how large they grow, they are usually not cancerous.

Revised Manuscript Received on December 30, 2019.
* Correspondence Author

Sarthak Garg*, Department of Computer Science and Engineering, Graphic Era deemed to be university, Dehradun, India. Email: garg.sarthak101@gmail.com

Harshit Garg, Department of Computer Science and Engineering, Graphic Era deemed to be university, Dehradun, India. Email: hgarg97@gmail.com

Vikas Tripathi, Department of Computer Science and Engineering, Graphic Era deemed to be university, Dehradun, India. Email: vikastripathi.be@gmail.com

Bhasker Pant, Department of Computer Science and Engineering, Graphic Era deemed to be university, Dehradun, India. Email: pantbhaskar2@gmail.com

Kumud Pant, Department of Computer Science and Engineering, Graphic Era deemed to be university, Dehradun, India. Email: pant.kumud@gmail.com

Although malignant skin cancer is the most dangerous among the type of skin cancers, almost all types of skin cancers are curable if detected at a early stage, enough for possible treatment. That's why it is very important to detect them as well as classify them based on category, as malignant cancers are very dangerous and need immediate attention. Malignant skin cancer if recognized early, can be easily treated and cured, otherwise it can become fatal and cancer can advance and spread adjoining cells, where it becomes hard to treat, In US annually around 9,320 people are killed by Skin Cancer. Out of which, 5,990 will be men and 3,330 will be women. This makes it important to discriminate between the benign and skin cancer. Deep learning is a class of machine learning methodologies used to solve demanding problems that arise in computer vision (CV) such as self-driving cars, digit recognition, image reconstruction etc., image processing such as pixel regeneration, colorizing black and white images, etc. Deep learning techniques such as Artificial Neural Network (ANN) and CNN give much more sophisticated results than the traditional machine learning techniques. Nowadays GPUs have high performance and computation power which makes the preparation of comprehensive datasets, less cumbersome which in turn gives results with better precision [2]. In the proposed paper, we have used a VGG-16 CNN framework which is pre-trained on the ImageNet Dataset and fine-tuned it to meet the requirements of the given classification problem [3].

The paper is comprised of following sections: 1- Introduction, 2- Literature Review, 3- Methodology, 4- Result and Discussion, and 5- Conclusion.

II. LITRATURE REVIEW

Formerly, researchers have proposed various techniques for classification of different types of cancer. Amirreza Mahbod et al. [4] used pre-trained VGG-16 and AlexNet models for feature extraction .These features are then utilized for training a non-linear multi-class Support Vector Machine (SVM) classifier for the evaluation of the classification results, then they used logistic regression to map SVM scores to probabilities, because the classifiers were trained for a multi-class classification into 3 classes namely seborrheic keratosis, melanoma and benign nevi and they achieved exceptional results in doing this. Xulei Yang et al. [5] have used Deep Convolution Neural Network (DCNN) to perform lesion segmentation and binary lesion classifications on Melanoma type skin cancer dataset from ISIC challenge 2017.

They utilized a multi-task DCNN strategy that enables distinctive segments to share the distinguished features among various attribute categories. These components This will allow hearty detection of the skin lesion from the dataset. Lei Bi et al. [6] have used a fine tuned ResNets Model pre-trained on ImageNet dataset for lesion segmentation classification of three classes on a similar dataset and achieved good results. They compared the proposed segmentation method (ResNet-Seg) with the FCN architecture (VGGNet model) and further evaluated the performance of using the additional training images and the proposed multi-scale integration approach. Noel C. F. Codella et al. [7] have paired original images with manual tracing of lesion boundaries to perform legion segmentation. Andre Esteva et al. [8] used CNN that was trained using 757 disease classes. They made use of GoogleNet Inception v3 model, a architecture pre-trained on around 1.28 million images from the 2014 ImageNet Large Scale Visual Recognition Challenge (ILSVRC), and implemented it on the dataset using transfer learning. Zongyuan Ge et al. [9] have used VGG-16 model comprising of 3 fully-connected and 13 convolutional layers, (VGG-BL). They used VGG-BL network which utilizes bilinear pooling (BP). The BP layer is infused with the non-linearity function (ReLu) after the last convolutional layer. The last VGG-16 convolutional layer has 512 feature kernels which leads to bilinear feature dimensions of 512x512, they used ResNet with the original pre-trained specifications taken from the ImageNet models and after the last ReLU layer inserted a bilinear pooling layer. Teck Yan Tan et al. [10] have used Generic Algorithm (GA) for optimizing the features and segregate most distinctive features for classification. This process removes inessential features thus improving accuracy. Later, the most distinctive features were used to classify between healthy and cancerous skin conditions by giving those features as input to SVM. Catarina Barata et al. [11] presented a different approach for melanoma identification that make use of dermoscopy imagery, for feature extraction based on texture of surface and color shades. They compared the performance on Texture and color basis, in which the color turned out to be more effective for isolation. Kiran Ramlakhan et. al. [12] developed a portable classifier application for epidermal lesion classification. They used the computer vision library of python named OpenCV for contour detection and image segmentation, later on feature extraction and classification is performed on the images.

III. METHODOLOGY

There are a number of proven techniques for image classification but Deep Learning techniques such as CNN have proven to be the most effective among the lot [13].
For robust skin cancer image classification, the brightness of the images has to be adjusted to obtain better features during the feature extraction stage. This brightness adjustment is done in the preprocessing stage of the proposed framework (Fig 1.). Other tasks done in this stage includes grayscale conversion and edge detection using canny method [13].
After the preprocessing stage, the pre-processed images are fed to a Fine-tuned VGG-16 Pretrained model.

generate features representations explicit to the attribute classes, and a multi-task training on the features can be used to deduce the attributes.

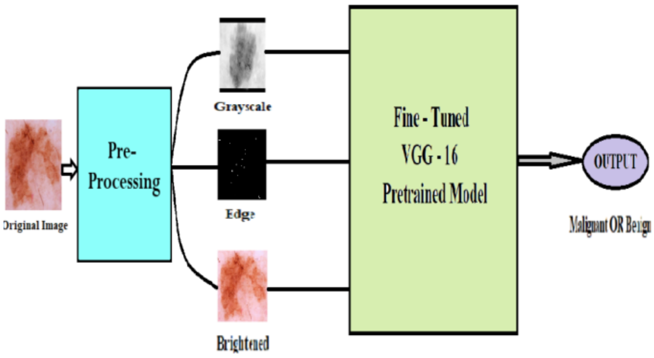


Fig 1: Proposed Approach Framework

VGG16 is a CNN model proposed in ILSVRC 2014 by K. Simonyan and A. Zisserman [14]. CNN is a class of neural nets that comprises of convolution layer, these layers execute convolution operations on the input and give output to the next layer [15], CNN uses the technique of backward propagation. The first convolutional network was presented in 1989 [16]. CNN comprises of 4 steps: 1-Convolution, 2-Max Pooling, 3-Flattening, 4-Full Connection.

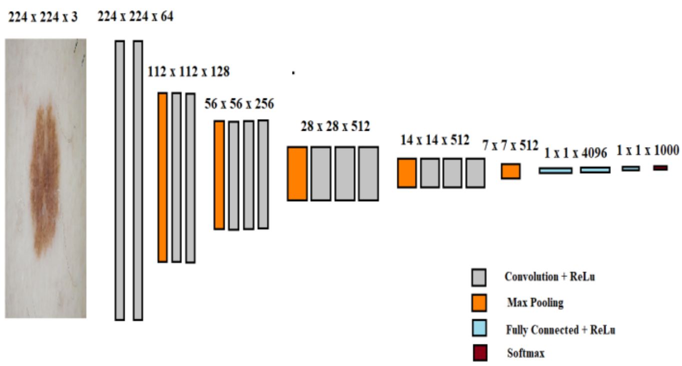


Fig 2: Pretrained VGG-16 Model

The pretrained VGG-16 model (Fig 2) achieved top-5 test accuracy of 92.7% in ImageNet, the dataset contained 1000 classes and around 14 million images. VGG-16 was one of the most renowned models proposed in ILSVRC 2014. A fixed RGB image of 224x224 is fed to the first convolution layer. The images are fed to a network of convolutional layers which uses very small filters of size 3x3. In one of the configurations, After the input channels are linearly transformed a ReLu unit is added with a 1x1 convolution filter. To preserve the spatial resolution after convolution, the convolution stride is fixated to 1 pixel. The small-size convolution filters allow VGG to have a huge number of weight layers; of course, more layers lead to improved performance. Spatial pooling is performed with the help of 5 max-pooling layers, some of the convolution layers are followed by these max pooling layers and this maxing polling is done by using 2x2 filter with a stride of 2.
VGG has 3 fully connected (FC) layers as demonstrated in (Fig. 2.),

The first and second FC layers both contain 4096 channels whereas there are only 1000 channels in the final FC layer, where each of the 1000 channels is representing one class in the classification.

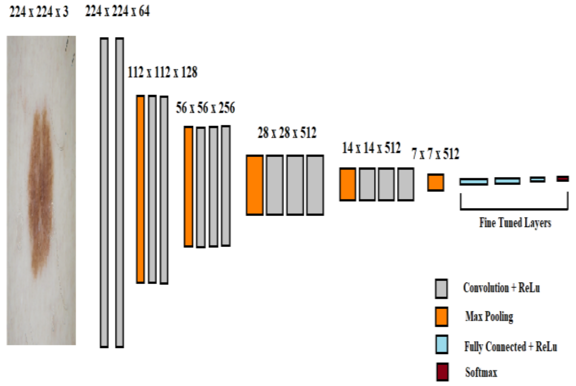


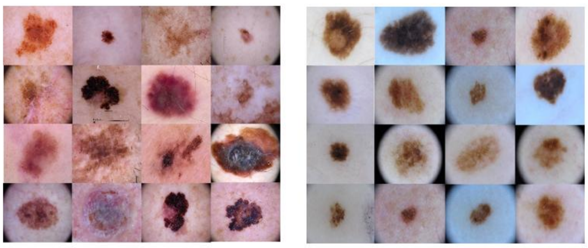
Fig 3: Fine Tuned Pretrained VGG-16 Model

In our proposed framework, the last 4 layers of the pretrained VGG-16 architecture are fine-tuned, and remaining layers which are responsible for extracting the dominant features from the input images are frozen. The pre-processed images are fed into this proposed VGG-16 Architecture in which, of the 4 fine-tuned layers the first 2 FC layers are made with 1024 number of channels each and then are further manipulated get the best results for the classification of our dataset into 2 classes.

The 3 FC layers follow a network of convolutional layers, the first 2 layers have been tried and tested with different number of channels, the third layer contains 2 layers which is used for 2-way classification. A soft-max layer is added at the end for probability distribution.

IV. RESULTS AND DISCUSSION

The approach proposed here uses a fine-tuned VGG-16 network to gather the dominant features needed to classify skin cancer into broadly two types. The proposed approach was implemented on a machine constituting of a 16 GB RAM, a GPU of 4GB Nvidia Gtx 1050Ti, and an Intel i7 7th Gen octa-core processor.



Malignant **Benign**

Fig 4: Sample Dataset

The deep learning classification models require a hefty dataset to obtain good testing results. The available images from ISIC archive [17] gallery was used for training purposes. The dataset used for this framework consists of 3801 colored dermoscopic skin images prorated into two classes, malignant and benign. There are 1896 malignant skin cancer images and 1905 benign skin cancer images in the dataset, the images at hand are of various dimensions (varying from 1022 × 767 to 6748 × 4499), lightening conditions and photographic angles. Some sample dataset images are shown in Fig 4.

Table I: Size of Dataset

Class	Number of Images
Malignant	1896
Benign	1905
Total	3801

Dataset shown in Table 1, is divided into 70% images in the training set and the rest 30% in test set for both classes.

All the images of a fixed size of 224x224 are fed to the preprocessing stage. In the pre-processing state edge detection was performed on the images, which is one of the most dominant features during the feature extraction stage, and carried out the classification on these images but got an accuracy of only 90.23 % which is not a satisfactory result. Thus, the original images are converted to grayscale, which has proven from time to time to be the widely used image type as it reduces computational time as well as give fair results but, this framework gave a subpar accuracy of 93.26% when fed with grayscale images. After getting underwhelming results when the framework was trained with the above two pre-processed images, finally we trained the framework on the images after brightness adjustment done in the pre-processing, though these images took the longest to be trained on our framework but it gave the best accuracy of a staggering 96.07% out of the lot.

Table II: Accuracy Assessment of different image types

Image Type	Accuracy
Edge Detected Images	90.23 %
Gray Scale Images	93.26 %
Brightened Colored Images	96.07 %

After getting the best accuracy in the Brightened coloured images we further continued our analysis by manipulating the channel count in the fine-tuned layers of our model. Firstly, we tried taking 2048 and 1024 as channel count in the first 2 FC layers, which yielded an accuracy of 96.98%, further we tried increasing the channel count of the 2 FC layers to 2048 and 2048, and achieved a staggering accuracy of 97.81%. Observing this exponential growth in accuracy on increasing number of channels, we increased the number of channels to 4096 and 2048, but contrary to expectations the accuracy dipped to 96.2%. On further increasing the number of channels to 4096 and 4096, the accuracy dipped to 94.8%. Observing the above pattern, it can be concluded that 2048 and 2048 number of channels yield best accuracy among the lot. Table 3 gives the complete accuracy assessment of different channel count in the fine-tuned layers of the Brightened colored images.

Table III: Accuracy Assessment of different number of channels in the fine-tuned layers

Number of channels in FC Layer 1	Number of channels in FC Layer 2	Accuracy
1024	1024	96.07 %
2048	1024	96.98 %
2048	2048	97.81 %
4096	2048	96.2 %
4096	4096	94.8 %

Graph in (Fig.5.), represents the accuracy variation with respect to the variation in the channel count in the first 2 FC layers of a pretrained VGG model.

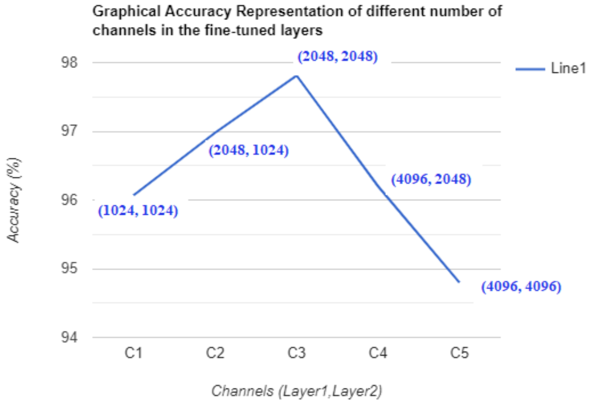


Fig 5: Accuracy Representation of different channel count in the fine-tuned layers

V. CONCLUSION

In this paper, we have applied a fine-tuned deep learning framework which uses the pretrained VGG-16 model to classify skin cancer into benign and malignant, and have obtained a competent accuracy of 97.81%. After testing out manipulations in various parameters such as image type, channel count in the first 2 FC layers of the VGG-16 framework we have concluded that our proposed framework model works best when the first 2 FC layers of the VGG network comprises of 2048 channels each and fed with brightened images of size 224x224. The model uses a compilation of nearly 4000 images of good quality with proper color and dimensions. These images include both the malignant and benign type of skin cancer. In future, there is scope of improvement to the proposed framework by further preprocessing of the images with advanced techniques as well as increasing the current dataset.

References

1. Aim at melanoma. (n.d.) Retrieved from www.aimatmelanoma.org: https://www.aimatmelanoma.org/about-melanoma/other-lesions/benign-or-malignant/
2. Krizhevsky, A. I. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 1097-1105
3. VGG16 – Convolutional Network for Classification and Detection. (n.d.) Retrieved from www.neurohive.io: https://neurohive.io/en/popular-networks/vgg16.
4. Amirreza M., Rupert E. (2017, Feb 27). Skin Lesion Classification Using Hybrid Deep Neural Networks. Published at Cornell University: https://arxiv.org/abs/1702.08434
5. Yang X., Zeng Z., Yeo S., Tan C., Tey H. & Su Y. (2017, March 17). A Novel Multi-task Deep Learning Model for Skin Lesion Segmentation and Classification. Published at Cornell University: https://arxiv.org/abs/1703.01025
6. Lei B., Jinman K., Euijoon A., & Dagan F. (2017, March 12). Automatic Skin Lesion Analysis using Large-scale Dermoscopy Images and Deep Residual Networks. Published at Cornell University: https://arxiv.org/abs/1703.04197
7. Codella N., Gutman D., Celebi M. et al. (2018, May 24). Skin lesion analysis toward melanoma detection: A challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). Published at IEEE, Washington DC: https://ieeexplore.ieee.org/abstract/document/8-363547
8. Esteva A., Kuprel B., Novoa R., Ko J. et al. (2017, February 2). Dermatologist-level classification of skin cancer with deep neural networks. Published at Nature International Journal of Science: https://www.nature.com/articles/nature21056?TB_iframe=true&width=914.4&height=921.6
9. Ge Z., Demyanov S., Bozorgtabar B., Abedini M. (2017, June 19). Exploiting local and generic features for accurate skin lesions classification using clinical and dermoscopy imaging. Published on IEEE: https://ieeexplore.ieee.org/document-/7950681
10. Tan T., Zhang L. & Jiang M. (2016, August 15).An intelligent decision support system for skin cancer detection from dermoscopic images. Published at 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD): https://ieeexplore.ieee.org/document/7603521
11. Barata, C., Ruela, M., Francisco, M. et. al. (2013, July 30). Two Systems for the Detection of Melanomas in Dermoscopy Images Using Texture and Color Features. Published at IEEE Systems Journal (Volume: 8 , Issue: 3, Sept. 2014): https://ieeexplore.ieee.org/abstract/document/6570764
12. Ramlakhan, K. & Shang, Y. (2011). A Mobile Automated Skin Lesion Classification System. Published at 2011 23rd IEEE International Conference on Tools with Artificial Intelligence: https://ieeexplore.ieee.org/document/6103318
13. Nijhawan, R., &Verma, R. (2017). An Integrated Deep Learning Framework Approach for Nail Disease Identification.2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), (pp. 197-200).
14. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
15. Canny J. (1986, November 6). A Computational approach to edge detection. images. Published at IEEE Transactions on Pattern analysis and Machine Intelligence: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.420.3300&rep=rep 1&type=pdf.
16. Wikipedia (n.d.) Convolutional neural network. Retrieved from: https://en.wikipedia.org/wiki/Convolutional_neural _network
17. ISIC Archive (n.d.) ISIC Archive Gallery.Retrieved from: https://www.isicarchive.com/#!/topWithHeader/onlyHeader Top/gallery

AUTHORS PROFILE

Sarthak Garg is pursuing B.Tech. in Computer Science and Engineering from Graphic Era deemed to be university, Dehradun, India. Interested in the field of Machine Learning, Deep Learning and Computer Vision, likes to work on innovative emerging technologies in the field of Computer Science to solve challenging problems.

Harshit Garg is pursuing B.Tech. in Computer Science and Engineering from Graphic Era deemed to be university, Dehradun, India. Data Science and Machine Learning Enthusiast, likes to work on challenging world problems and develop new software for the advancement of the society.

Dr. Vikas Tripathi has done BE in information technology from Technocrats institute of technology, Bhopal, M. Tech in Software engineering from Indian institute of information technology Gwalior and PhD from Uttarakhand technical university, Dehradun. He is actively involved in research related to Software engineering, Computer Vision, Machine learning and Video Analytics. He has published many papers in reputed international conferences and journals. Currently he is working as an associate professor in Graphic era deemed to be university Dehradun, India.

Dr. Bhasker Pant Currently working as Dean Research & Development and Associate Professor in Department of Computer Science and Engineering. He is Ph.D. in Machine Learning and Bioinformatics from MANIT, Bhopal. Has more than 15 years of experience in Research and Academics. He has till now guided as Supervisor 3 Ph.D. candidates (Awarded).and 5 candidates are in advance state of work. He has also guided 28 MTech. Students for dissertation. He has also supervised 2 foreign students for internship. Dr.Bhasker Pant has more than 70 research publication in National and international Journals. He has also chaired a session in Robust Classification & Predictive Modelling for classification held at Huangshi, China.

Kumud Pant has been awarded PhD in Bioinformatics from MANIT, Bhopal in 2014. Masters in Biotechnology from Jiwaji University, Gwalior. Granted full air fare from CSIR for attending IC2IT held at Bangkok, Thailand in 2009. Project from UCOST entitled "Reverse vaccinology approach for detecting major virulent proteins of Encephalitis virus". Ph.D. supervisor for two (2) and co-supervisor for one (1) student. One (1) student Mr. Vipin Sati awarded Ph.D. (Co-Supervisor).