International Skin Imaging Collaboration: Advancing Melanoma Detection through Digital Skin Imaging and Standardization

Divya Athapuram[1], Kattipally Radha Reddy[2], Priya Jayaraman[3]

[1, 2,3]Assistant Professor, Department of Information Technology, Malla Reddy Engineering College and Management Sciences, Hyderabad, Telangana.

Abstract

The International Skin Imaging Collaboration: Melanoma Project is a collaborative initiative between academia and industry aimed at harnessing the potential of digital skin imaging to reduce melanoma-related mortality. Early detection and treatment significantly improve the prognosis for melanoma, making it a highly curable disease when identified in its initial stages. Digital skin lesion images can play a vital role in educating both medical professionals and the public in recognizing melanoma while also directly assisting in melanoma diagnosis through practices like teledermatology, clinical decision support, and automated diagnosis. However, the absence of standardized practices in dermatologic imaging has been a hindrance to the quality and utility of skin lesion imaging. To address these concerns, the ISIC project is actively working to develop and propose standards encompassing the technologies, methodologies, and terminology used in skin imaging. This endeavor pays particular attention to matters of privacy and interoperability, ensuring that images can be securely and seamlessly shared across diverse technological and clinical platforms. Furthermore, ISIC is establishing an open-source public archive of skin images, which serves as an invaluable resource for education and the development and validation of automated diagnostic systems.

Keywords: Melanoma Detection, Skin Imaging, Standardization, Teledermatology, Clinical Decision Support, Open-Source Archive, Dermatologic Imaging.

1 Introduction

Cancer nowadays is one of the greatest growing groups of diseases throughout the world, among which skin cancer is most common of them. According to stats and figures, the annual rate of skin cancer is increasing at an alarming rate each year [1]. The modern medical science and treatment procedures prove that if skin cancer is detected in its initial phase then it is treatable by using appropriate medical measures which includes laser surgery or removing that part of the skin which ultimately could save a patient's life. Skin cancer has two main stages which include malignancy and melanoma among which melanoma is fatal and comes with the highest risk. In most cases, malignant mole is clearly visible on the patient's skin which is often identified by the patients themselves. Dermoscopic diagnosis refers to a non-invasive skin imaging method, which has become a core tool in the diagnosis of melanoma and other pigmented skin lesions. However, performing dermoscopy using conventional methods may lower down the diagnostic accuracy which can lead to more chances of errors. These errors are generally caused by the complexity of lesion structures and the subjectivity of visual interpretations [2]. Computer-Aided

Diagnosis (CAD) system is a type of digitized platform based on advanced computer vision, deep learning, and pattern recognition techniques for skin cancer classification. For the proposed study we have designed a CAD system for skin cancer classification by utilizing advanced deep neural networks. The system consists of the following steps: Firstly, a preprocessing of the digital images which includes removing clutter such as hair from that part of the skin where the pigmented mole is present and applying a sharpening filter to make that area more clear and visible thus minimizing the chances of error.

2 RELATED WORK

In dermatology, there has been a long interest in exploiting computer technology. Recent years have seen increased activities in developing in machine learning and computer vision techniques for skin lesion diagnosis, especially for diagnosing melanoma case. Braun et al. have worked on [1] dermoscopy research tool and also cover different aspects, such as the new equipment, new structures, the importance of blood vessels, etc. There were several drawbacks in dermoscopy; it could not be used for clinically suspicious skin lesions. Silveira et al. proposed the early diagnosis of malignant melanoma, but their interpretation is time consuming and subjective, even for trained dermatologists[2].Six different segmentation methods are Adaptive thresholding , Gradient vector flow, Adaptive snake, Level set method of Chan et.al, Expectation-maximization level set , Fuzzy-based split-and-merge algorithm were compared and evaluated by four metrics.(HM,TDR,FDR,HD).Out of six segmentation methods, only AS and EM-LS methods are robust and useful for the lesion segmentation to assist the clinical diagnosis of dermatologists. Rademaker et al. [3] introduced digital monitoring by whole body photography and sequential digital dermoscopy detects thinner melanomas. Patients undergoing whole-body photography and sequential digital dermoscopy are largely self-referred. Melanoma not is detected until they are quite advanced. Abbasa and Celebic introduced a comparative study of the stateof-the-art hair-repaired methods with a novel algorithm is also proposed by morphological and fast marching schemes [4]. Nonlinear partial differential equation based methods are not texturebased inpainting methods and, it was not suitable for hair removal in dermoscopic images. Suer et al. [5] developed a automated assessment tools for dermoscopy images have become an important research field mainly because of inter- and intra-observer variations in human interpretation. It works on color image without preprocessing and it cannot find any point density-reachable from the starting point. This procedure followed until all of the points in the EPS neighborhood are touched or visited at least once.

3. PROPOSED METHOD:

3.1. Convolutional Neural Networks

The basic purpose of pooling in CNN is the task of subsampling i.e., it summarizes the nearby neighborhood pixels and replaces them in the output at a location with summarized characteristics. Pooling reduces the dimensionality and performs the invariance of rotational transformations and translation transformations.

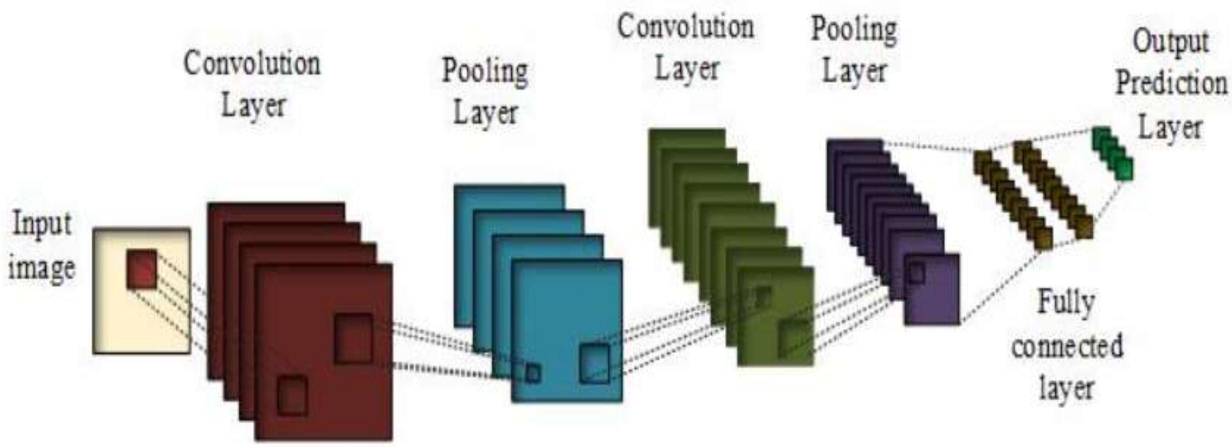

Figure 1. General convolutional neural networks

There are many pooling functions [7]; one of the most famous is max pooling, in which the output is the maximum value of the rectangular pixel neighborhood. In average pooling function, the output becomes the average of the rectangular neighborhood. Another type consists of the weighted average based on the distance from the central pixel. Pooling helps to make the representation invariant to small changes to the translation in the input. Deep residual learning is used to counter the degradation problem, which arises when the deep network starts to converge, i.e., a saturation of accuracy and degradation with the increasing depth. The residual network explicitly allows the stacked layers to fit in the residual map rather than a desired underlying map. According to the experimental results, the optimization of residual networks is easier, and the accuracy is achievable with a considerable increase in depth. Skip connections help the transverse information in deep neural networks. Due to passing through many layers, the gradient information may be lost, which is known as the vanishing gradients problem. Skip connection has the advantage of passing the feature information to lower layers, which makes it easier to classify the minute details. Some of the spatial information is lost due to the max-pooling operation, whereas skip connections make it possible to have more information on the final layer so that the classification accuracy increases.

Initially, researchers invented an algorithm, namely “hole algorithm” or “algorithme à tours” for wavelet transformation [8], but right now in the deep learning area is known as “atrous convolutional” or, “Dilated Convolution”. The dilated convolution expands the kernel’s field of view with the same computational complexities by insert “hole” or zeros between the kernel of each convolutional layer. Therefore, it can use for those applications which cannot bear bigger kernels or, many convolutions, however, require a wide field of vision.

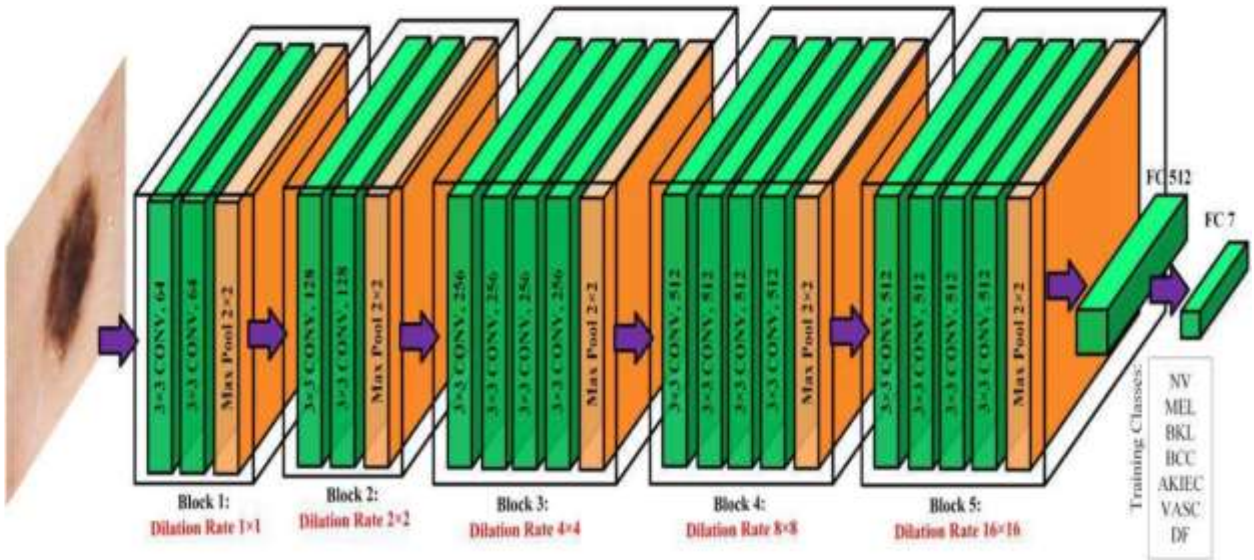

Fig 2: proposed deep learning architecture

Both VGG16 and VGG19 have five blocks of convolutional layers were with an equal number of parameters to expand the context view of filters where we modify the dilation rate of these layers. Output feature map will shrink (output stride increasing) for any standard convolution and pooling if we go deeper in any model, which is harmful for classification because, in the deep layers, spatial information will be missing. With dilated convolution without increase computational complexities, we can achieve a larger output feature map, which is proved to be appropriate for skin lesion classification in terms of accuracy. Both the model has 3 × 3 kernel size in every layer, and without increasing kernel size, we can enhance the receptive field dimension by adding different dilation rates in the existing layer in our proposed architectures. The input layer size is 192 × 256 × 3 ($height\ of\ the\ image \times width\ of\ the\ image \times RGB$). The initial block of these networks have $dilation\ rate = 1$, then from block two to five have dilation rate 2, 4, 8, and 16 respectively. To implement this technique, we mostly inspired by different multi-grid models where we find a hierarchy of several different sizes of grid [53-56], many semantic image segmentation models [57, 58], in [58] Chen et al. pick distinct dilation rates within block4 to block7 in the proposed ResNet model. Seemingly, we utilize VGG networks to adopt different dilation rates in different blocks. In the all convolutional layer, we used rectified linear units (ReLUs) as activation function and max-pooling used for downsampling in between every convolutional block. After the last convolutional and max-pooling layer, we run global max-pooling operation which takes tensor with shape $h \times w \times d$ ($h \times w$ = spatial dimensions, $d = number\ of\ feature\ maps$) and provides output tensor with shape $1 \times 1 \times d$. Then, we add two fully connected layers in these models with 512, 7 (dataset has seven classes) filters respectively with a dropout layer ($dropout\ rate = 0.50$) in between which utilize as a regularizer function to substantially weaken the overfitting rate and computationally reasonable at the same time [59]. "RELU" is the activation function for the first dense layer, and the last one has "SoftMax". From Figures 1 and 2, we can notice the details visualization of dilated VGG16 and VGG19. MobileNet was constructed to provide small, very low latency, and computationally sound model for embedded mobile vision applications [47]. MobileNet has three kinds of convolutional operation: standard convolution, pointwise convolution, and depthwise convolution. We take five depthwise layers to implement the dilated convolution, and every layer has a stride rate (2,2). Among these depthwise layers, the first two layers have a dilation rate (1,1); however, for the third and fourth layers, we placed a dilation rate (2,2). Furthermore, for the final depthwise layer, we concatenate three depthwise 2D convolution layers parallelly with a dilation rate of 4,8,16, respectively. Finally, we concatenate these three 2D layers and produce the fifth depthwise convolutional 2D layer. Originally, every depthwise layer of MobileNet has $dilation\ rate$ = 1, but implementing different dilation rates in distinct depthwise layers of MobileNet architecture is new, and we first propose this approach.

After all the convolution operation (standard, depthwise, and pointwise), from the last pointwise layer, we take the feature map and employ the global average pooling (GAP) method. Global average pooling converts the feature map size into $1 \times 1 \times d$ from $h \times w \times d$, and here this method takes average value from the spatial dimension of the feature map ($h \times w = spatial\ dimension$). GAP has several advantages; such as elude overfitting in the layer, in the input feature map, it exhibits more robust characteristics to the spatial translations [60]. The classifier part of the fully connected part is the same as the VGG networks. There are two fully connected layers, and in between, there is a dropout layer.

4. EXPERMENTAL RESULTS

Next, Pandas and Scikit Learn utilized respectively for data preprocessing, and to evaluate these proposed models. Training every model for 200 iterations and take 32 as the mini-batch size. The models executed on Intel Core i7-8750H with 4.1 GHz and an NVIDIA GeForce GTX 1050Ti GPU. Adam 64 optimizer used as the optimization function with a learning rate 10-4 initially. One callback function utilized to lessen the learning rate factor by (0.1).5 during the training when the loss of validation is not diminishing for seven iterations. Thus, the new learning rate: $New\_lr = lr * (0.1).5$ (2) New_lr = new learning rate; lr = present learning rate. Here, the lower bound for the overall learning rate is $0.5e - 6$.

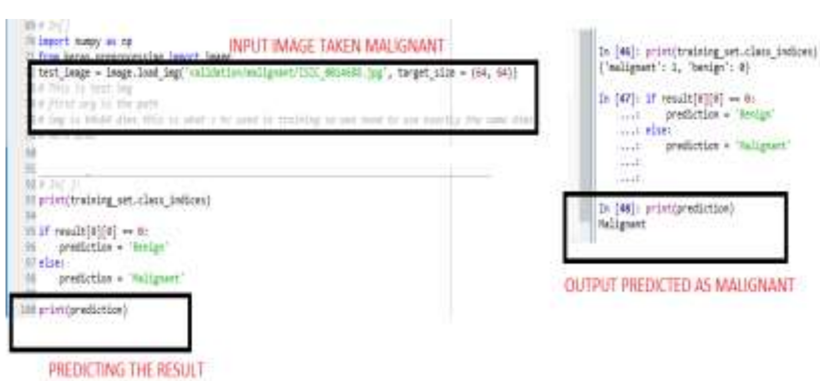

Fig 3: Prediction result for malignant skin cancer

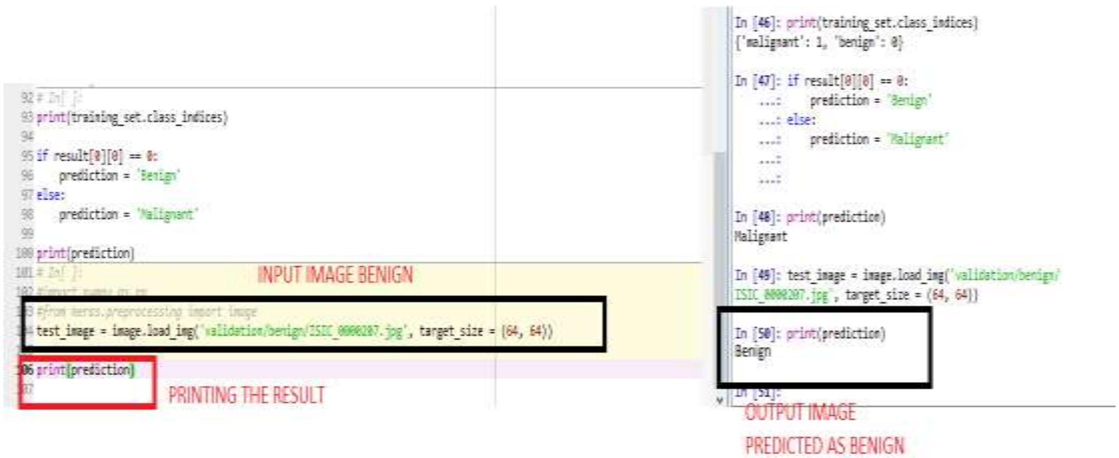

Fig 4: Prediction result for benign skin cancer

Top-1 Test Accuracy of Four Models Before proposed these models, we tried numerous combinations for these four architectures. After the experiment with several combinations, we able to fixed which design for each model produces the best top1-accuracy and per-class accuracy. Furthermore, we examine different image resolution (64×64,128×96, 256×192, 320×320) for proposed VGG16, VGG19, and InceptionV3. Overall, these three networks produce the best outcome for 256×192 image resolution. On the other hand, among the different image size combinations for dilated MobileNet (128×128, 160×160, 192×192, and 224×224) and 224×224 image shape provide the highest accuracy. From table 2, we can see the dilated InceptionV3 showed the foremost top-1 accuracy among these four models, and it displayed superior computational complexities. However, dilated MobileNet provides lightest computational complexities with only 3.7 million parameters.

Fig. 3,4 shows some of the results predicted correctly on test images. Overall, in both networks, significant improvements were measured after using the refined version of images. The experimental results show that the Inception-v3 network was able to achieve better validation accuracy using a refined version of training data i.e. 86.1 % thus we will be using the Inception-v3 network for evaluating it on the test data. For evaluating the classifiers on the test data, we have picked numerous cases from the test set from both classes, benign and malignant melanoma among which visually complex and challenging test cases were selected for the proposed research work. It is pertinent to mention that the network was tested using the original images (unrefined version) to test the overall effectiveness of the classifier.

5. CONCLUSION

Skin cancer is one of the dangerous forms of cancer as the affected cells can spread easily across the body. It can be either Melanoma or Non-Melanoma. There are various solutions such as Ceroscopy and other devices to detect the skin cancer, these devices involve costs as well as requires a doctor to equip them on the patients. Proposed method aims at detecting and prediction of skin cancer using Image Processing Techniques that can be easily used by Doctors for the Patient's skin cancer analysis. The system employs methods such as Preprocessing, Feature Selection, Feature Extraction and Back propagation Neural Networks. The outcome of the model is determined by the BPN (Back propagation Net-work) that predicts the type of the cancer. This kind of models helps the patients to take care of their skin as well as take precautionary measures if the Skin cancer is encountered. The model was applied on a ISIC (International Skin Imaging Collaboration) dataset and resulted in the classification of the cancer types.

REFERENCES

[1] J. R. P. Braun, H. Rabinovitz, J. E. Tzu, and A. A. Marghoob, ``Dermoscopy research An update," Seminars Cutaneous Med. Surgery, vol. 28, no. 3, pp. 165-171, 2009.

[2] M. Silveira et al., ``Comparison of segmentation methods for melanoma diagnosis in dermoscopy images," IEEE J. Sel. Topics Signal Process.,vol. 3, no. 1, pp. 35-45, Feb. 2009.

[3] M. Rademaker and A. Oakley, ``Digital monitoring by whole body photography and sequential digital dermoscopy detects thinner melanomas,"J. Primary Health Care, vol. 2, no. 4, pp. 268-272, 2010.

[4] Q. Abbas, M. E. Celebi, and I. F. García, ``Hair removal methods: A comparative study for dermoscopy images," Biomed. Signal Process. Control,vol. 6, no. 4, pp. 395-404, 2011.

[5] S. Suer, S. Kockara, and M. Mete,"An improved border detection in dermoscopy images for density based clustering," BMC Bioinformat., vol. 12, no. 10, p. S12, 2011.

[6] T. Wadhawan, N. Situ, K. Lancaster, X. Yuan, and G. Zouridakis, "SkinScan: A portable library for melanoma detection on handheld devices," in Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro,Mar./Apr. 2011, pp. 133-136.

[7] C. Doukas, P. Stagkopoulos, C. T. Kiranoudis, and I. Maglogiannis, "Automated skin lesion assessment using mobile technologies and cloud platforms," in Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Aug./Sep. 2012, pp. 2444-2447.

[8] M.Emre, CelebiQuan, WenSae, HwangHito, shiIyatomi GeraldSchaefer"Lesion Border Detection in Dermoscopy Images Using Ensembles of Thresholding Methods",Dec 2013.

[9] Q. Abbas, I. F. Garcia, M. E. Celebi, andW. Ahmad, "A featurepreserving hair removal algorithm for dermoscopy images," Skin Res. Technol.,vol. 19, no.1,2013.

[10] O. Abuzaghleh, B. D. Barkana, and M. Faezipour, ``Automated skin lesion analysis based on color and shape geometry feature set for melanoma early detection and prevention," in Proc. IEEE Long Island Syst., Appl. Technol. Conf. (LISAT), May 2014.

[11] A. Karargyris, O. Karargyris, and A. Pantelopoulos, ``DERMA/Care: An advanced image-processing mobile application for monitoring skin cancer," in Proc. IEEE 24th Int. Conf. Tools Artif. Intell. (ICTAI),Nov. 2012.

[12] O. Abuzaghleh, B. D. Barkana, and M. Faezipour, "Noninvasive real time automated skin lesion analysis system for melanoma early detection and prevention" in Proc. IEEE., Apr. 2015.