Malignant melanoma classification using machine learning algorithms.

Rashmi Patil [*]**, Sreepathi Bellary**

Department of Computer Science and Engineering, Rao Bahadur Y Mahabaleswarappa Engineering College, Ballari, Karnataka, India

Abstract

Nowadays growth of malignant skin causes the death of patients and is a highly identified reason for death among people. The unusual growth of skin cells on any part of the body happens, and it comes to exposure of sunlight is known as malignant skin. If malignant skin cancer is detected at an early stage, then a huge portion of skin malignant growth is recoverable so it can save the patient's life. The identification of melanoma skin cancer is possible at an early stage by using novel methods. Here we described two schemes for detecting melanoma from benign dermoscopic pictures. The first scheme combines the Convolutional Neural Network (CNN) with the Support Vector Machine (SVM), that is, CNN+SVM and the second scheme combines CNN with Extreme Gradient Boosting (XGB), that is, CNN+XGB, because CNN works well for the training dataset, and the SVM work fine in classification also XGB used as a recognizer. The total of 2000 images are taken and achieved accuracy 84.0% for CNN+SVM and for 88.0% CNN+XGB.

Introduction

Malignant melanoma, is referred as a "melanoma," and is a kind of skin cancer that in almost all cases begins in pigment cells of the skin. Almost all people when they think about melanoma think of a dark grown lesion; after all, these tumors grow from pigment cells. The skin is basic part of body ensures against daylight, heat, some destructive beams from the sun, various diseases and wounds. Skin is performing significant job to store water, nutrient D and fat and keep up internal heat level. Malignant growth is dangerous type of skin cancer [1,2]. Malignant growth occurs as unrepairable damage to the skin's cells, is in part caused by sunlight, and inherited malformations. Melanoma looks similar to moles or also might be generated from moles. The color of melanoma might be dark, brown, purple, red, skin-hued, pink, white or blue. Few melanomas lose pigment, without dark pigment partially or fully, and become pink, white, or tan. Apart from color, elevation, there are numerous other features of melanomas like texture [3] and an existence of certain structures in the lesion that separate them from benign lesions. If detection of melanoma happens at the beginning stage then it can be treated and cured. The malignant growth happens on numerous body parts, and it leads to difficulties to cure. Dermoscopy [4] or microscopy of the dermoscopy is often referred to as dermatoscopy. Dermoscopy is a procedure for magnifying and shining pictures of a section of the skin to enhance the integrity of the skin spots. The device associated with the image is known as a dermatoscope. Touch, using a gel coating between the skin, dermatoscope, and contact-free, without skin or fluid contact are known as dermatoscopes. Cross polarized light from the dermatoscope is used for the picture in non-contact pictures and some contact shots. Dermoscopy images are extremely used when skin lesions are recognized due to their enlightenment and magnification. Melanoma thickness [2] of melanoma needs to be detected for phases. Skin lesion detection [1] from dermoscopic pictures is a very difficult work for proficient dermatologists because of less difference among encompassing lesions and skin, the visual likeness between skin lesions, confused borders of lesions, and so forth. An automated dermoscopic [4,5] image detection device for the computer helps clinicians predict early damage to the malignant skin. In a wide sense, the profound learning is advanced and knowledgeable in terms of precision and comparable compatibility of the machine compared to CNN. To allow effective drugs, a blurred skin lesion needs to be identified. Early identification of melanoma [6] considerably rises survival rate of the patient. Only experienced dermatologists are able to detect exact melanoma skin lesions by seeing lesion dermoscopic images. Although, the detection of melanoma exactly is an incredibly difficult task due to the low difference between skin and lesion. So the non-invasive method is required to detect melanoma and benign. Proposed a novel system to reduce these challenges with difficulties for accuracy and automatic melanoma recognition [4] in dermoscopy images. Designed the framework for melanoma detection from benign by dermoscopic images. The first technique combines CNN with SVM, CNN [7] is utilized for training and SVM utilized for classification. The second scheme combines CNN with XGB, XGB utilized as recognizer. CNN is one of the classical deep learning [1] models. The human success in object recognition has been shown to go beyond that. Recently, photos consisting of dermoscopic skin cancers and artificial intelligence were recognized and demonstrated to be able to categorize skin cancer of a degree of

skin competence equivalent to dermatologist. In order to get better results, CNN is paired with SVM and XGB.

Literature Review

Here [8], proposed a Customized Convolutional Neural Network (CCNN) for analyzing handwritten Marathi numerals. The proposed model has an ability to work with large dataset and it can enhance the performance. In addition, it requires less processing time that is training time because of utilization of GPU and CPU. Also, the authors utilized ten-fold cross-validation techniques for better outcome. They utilized 80000 Marathi numerals as a dataset to train and test the results.

Melanoma is known as deadliest skin cancer because it causes person death. When it is diagnosed at the early stage, then it is recoverable so it is vital to classify them in stages, but it is a very complicated task to identify and classify them into stages. Here, [9] presented two methods that identify stages of melanoma. The first method classifies them into stage 1 and 2 and second method classify melanoma in stage 1, 2 or 3. Author utilizes CNN algorithm with a loss function Similarity Measure for Text Processing.

Here [7], presented novel melanoma classification method using CNN. This method utilized to classify dermoscopy images. They presented method known as a Region Average Pooling (RAPooling) is focuses the region of interest. Then introduced combination of classification with segmentation information which utilizes region of segmented lesion to direct the classification by RAPooling. Also introduced linear classifier Rank Opt to obtain optimized outcomes. The [10] introduced an approach by combining Conventional Image Processing (CIP) with deep learning. CIP utilized 3 handmade image processing modules and 1 clinical information module. In the image processing module detected lesion features as compared to clinical dermoscopy info. The clinical dermoscopy information consists of data given to clinical dermoscopy information and this module uses knowledge transfer through a ResNet-50 network. Handmade feature detection consists of the following steps: the first one is preprocessing then applied Median Color Split Algorithm and Telangiectasia Vessel Detection Algorithm Then Atypical Pigment Network Detection, Demographic Features detection, Salient Point Detection, Specific Colors Detection. Developed two Deep Learning (DL) methods for locating the three tasks arrived in the skin lesion image processing area [1]. These tasks are as follows lesion segmentation, dermoscopic feature extraction, and classification. The DL framework proposed to generate the segmentation outcomes and coarse classification results, the framework consists of two Fully Convolutional Residual Networks (FCRN). A Lesion Index Calculation Unit (LICU) is proposed to clarify the coarse classification results by estimating the distance heat-map. A straight-forward CNN is developed for the dermoscopic feature extraction task. DL framework uses ISIC 2017 dataset to estimate results. Here [2] presented a review on melanoma recognition and classification based on the thickness. Introduced numerous stages of melanoma. According to a review if cancer is diagnosed and treated at an early stage the survival chances of the patient will

be more. These stages depend on the thickness of melanoma hence we can say that survival chances are based on the thickness. And the thickness prediction of melanoma is very difficult there are very few techniques available for that. The [11] presented a competitive analysis for recognizing melanoma from dermoscopy images. Here the author presents a comparative study on two techniques of melanoma detection. The first method utilizes global features and the second method utilizes local features. Recognition is based on dermoscopy images. The detection based on the global feature method provided with dermoscopy images then performed segmentation of lesion then extracted features and applied classification on these features. And the detection based on the local feature method system provided with dermoscopy images of skin lesion then segments the lesion and selects the key points and then features are extracted based on that histogram drawn and performed classification. According to the authors, the local method is better than the global and the global strategy is better if the skin lesion is homogeneous. The [12] studied melanoma identification using deep learning CNN. Deep learning CNN demonstrates high potential for melanoma detection. Recognition of melanoma thickness based on position and subtype of melanoma. There was no question whether CNN can offset the diagnostic problems of the doctors in these melanomas. The focus of the Author is to research the diagnostics success of the CNN with European market approval in various localizations and subtypes of melanoma. A deeper regional network of convolutions and fuzzy mean C clusters used [13] to detect melanoma injuries and segmentation. Several areas of infection were accurately diagnosed by a profound, Regionally Focused Neural Network (RCNN) in the form of bounding boxes to simplify the location of the cluster. In order to efficiently, accurately and automated segment the Melanoma area into dermoscopic images, the author developed a new technique using RCMN and FCM clustering. The treatment is a three-stage procedure: skin flushing, melanoma area position and finally melanoma segmentation. The Deep Residual Networks for Automatic Dermoscopic Melanoma Detection, introduced very by [4]. Two measures, such as segmentation and classification, are the new approach proposed. The author integrates these two steps in full and produces a manually interactive and integrated system. Very deep CNNs could produce functions that are highly discriminatory as compared with shallow equivalents, both segmentation and gradation efficiency being improved. The author has also developed a new FCRN programmed for specifically segmented skin injuries. Here [5] presented a deep learning-based technique for detecting automatic melanoma, introduced to mitigate the limitations of existing system. For efficient learning and extracting the features from image an enhanced encoder-decoder network utilized. The proposed system consists of multi-stage and multi-scale approach with softmax classifier utilized to classify melanoma lesions pixel wise. The author designed lesion classifier method to do classification of skin lesions in non-melanoma and melanoma skin, for do that used outcomes obtained from pixel-wise classification. The main objective of [14] is to enhance skin lesion matching in false positive and negative form. So,

authors proposed two new strategies, first one is depending on successive rigid transformations of 3-D point clouds, and second is depends on non-rigid coordinate plane deformations. In both methods author applied a robust outlier rejection method, which relies on progressive graph matching. Author discussed about total body skin scanner. Here Total Body Photographs (TBP) are utilized, this TBP technique collect photographs of patients periodically and used by clinical experts to detect the melanoma. The early detection of melanoma is vital to save life of patient but all time it is not recognized exactly based on clinical appearance so additional methods needs to diagnose melanoma of pigmented skin lesions. Here [15] proposed *vivo* epiluminescence microscopy for detecting melanoma by detail analysis of pigmented skin lesion surface. Author utilize the oil immersion technique, provides the epidermis translucent, and opens a new dimension of skin morphology by including the dermoepidermal junction into the macroscopic evaluation of a lesion. They analyzed 3000+pigmented skin lesions using epiluminescence microscopy. Here [16] presented a phrase-based document similarity for estimating pairwise similarities of documents depend on the Suffix Tree Document (STD) model. phrase-based document similarity applied by author to group-average Hierarchical Agglomerative Algorithm of clustering (HAC) and definition of a new process called the clustering of documents. Suffix STD model node tree mapped into a single functional word with the Vector Space Document (VSD) model and a phrase document-oriented similarity tf-idf weight scheme used to estimate the document similitude [17] locate the two systems to identify melanoma utilizing dermoscopy images. In first system for classifying the skin lesions global methods utilized and in second method for classification local features and bag-of-features utilized. Here the main focus of the authors is to find best system among both. The novel approach proposed by [18] relies on a deformable paradigm for the treatment of dermoscopic skin lesion segmentations. Color RGB space to separate skin and skin lesions from normal. The proposed solution also provides powerful and versatile segmentation, vigorous against noise. For device performance monitoring, the author used two image databases and [6] presented 2 components for early identification and prevention of melanoma of a noninvasive real-time automated skin lesion analysis system.

These components are real-time alert and automated image analysis module. The first one helps people to prevent their skin from sunlight and second one consist of image property, hair diagnosis, segmentation of lesion, extraction of features and classifications. Author used PH2 Dermoscopy image database for testing performance of system and this database is taken from Pedro Hispano Hospital.

Proposed system

The proposed system Figure 1 for melanoma dataset classification and is a combination of CNN+SVM and CNN +XGB. A system is trained using CNN and activation function used is leaky relu.

Figure 2 shows proposed architecture in detail. The CNN+SVM and CNN+XGBoost architecture are constructed using SVM and XGBoost to create the last outcome layer for CNN frames. The Input Activation Function is a linear setting of previous hidden layer outputs, with qualified weights, as well as a biased time.

The hidden layer calculations are not only meaningful to the CNN framework, but can also be treated as input highlights for different classifications.

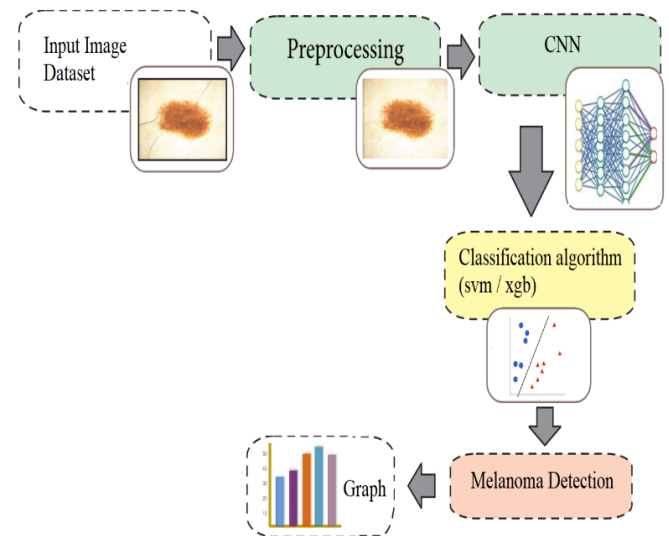The framework of this combination is focused on the advantages of CNN, SVM and XGBoost being consolidated.

Figure 1. Proposed System Architecture.

Classification using SVM and XGBoost

Supporting vector systems are supervised learning models with similar learning algorithms which define the data used for regression and assessment.

An SVM model portrays instances as space points mapped with the distance needed to categorize as much as possible the instances of the different categories. In the same space, new instances are then mapped and predicted to be part of a remote population.

The data collection of the image is processed and forwarded to the CNN input layer. The last CNN layer weights will be stored in the .h5 file after the planning. The SVM and XGB graders indicate these weights.

Here, the output layer of CNN is replaced by SVM and XGBoost, the CNN classifier soft-max, and the ability for training is used.
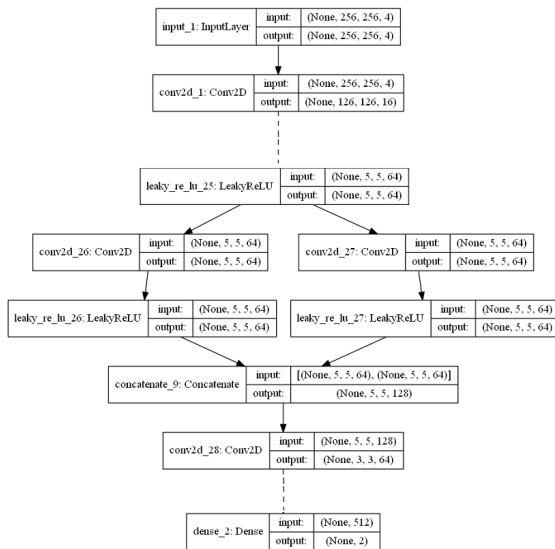
***Figure 2.** CNN archsitecture.*

Dataset description

We take dataset from ISIC 2017: Skin Lesion Analysis towards Melanoma Detection. This consists of training, validation and blind held-out test dataset. In this, we utilized ISIC dataset which consists, training: 2000 images testing: 600 images and validation: 150 images. Table 1 shows detailed about dataset: Class 0-Benign and Class 1-Melanoma.

	Total Images	Class 0	Class 1
Training dataset	2000	1626	374
Testing dataset	600	483	117
Validation dataset	150	120	30

***Table 1.** Dataset information.*

Preprocessing

The image dataset is preprocessed by utilizing numerous preprocessing techniques which consist of resize, segmentation, sharpening and reshape. Figure 3 shows preprocessed image.
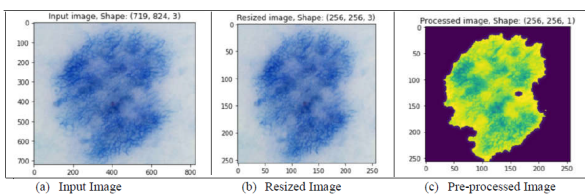

(a) Input Image   (b) Resized Image   (c) Pre-processed Image

***Figure 3.** Image pre-processing.*

Model training using CNN

A CNN Deep Learn algorithm can give significance to various aspects of the image (learnable weights and preference) and can distinguish between each aspect and the next. In contrast to other classification algorithms, ConvNet's preprocessing needs are significantly lower. CNN models will pass through a sequence of philter convolution layers, pooling, Full-Connected (FC) layers for training and testing, each input image and use Softmax/Sigmoid to classify an entity with probabilistic values between 0 and 1.

Proposed algorithm

Following are the steps involves in execution of our proposed system.

- ISIC images dataset: Input ISIC images dataset of the cancer infected skin (malignant and benign). Then
- Preprocessing: Perform preprocessing of the images: In this, the images will be resized, segmented, sharpened and reshaped.
- CNN: Apply CNN for training.
- The weights of last layer from CNN are stored in .h5 file which will be used by SVM and XGBoost classifiers as input.
- The CNN+SVM and CNN+XGB frameworks architecture are designed by substituting the last return layer for the CNN frames with SVM and XGBoost. In the last CNN laying, SVM and XGBoost are used to build new frameworks instead of the standard Softmax function with cross entropy.

Results and Discussion

A total of 100 children were enrolled in this study. Out of these 100 participants, 47 were males while 53 were females.

The age of the patients in this study ranged from 7 to 15 years with the mean of 12.07 ± 2.46 SD years as shown in Figures 1 and 3. 100 right eyes and 100 left eyes were analyzed.

Experimental setup

The proposed CNN+SVM and CNN+XGB systems are implemented with Python software and algorithms, and strategies congested with Anaconda (Jupiter), the opposing classification approach, and feature extraction techniques, and run in an environment with Intel Core i5-6200U, 2.30 GHz Windows 10 (64 bit) 8GB RAM configuration systems.

Graphs and tables

Figure 4 shows the graph of sensitivity and specificity comparison of CNN, CNN+SVM and CNN+XGB methods.

Figure 5 demonstrates the Precision, Recall, F-measure and Accuracy graph comparison of CNN, CNN+SVM and novel CNN+XGB methods. The outcomes demonstrate CNN+XGB gives more accurate results.
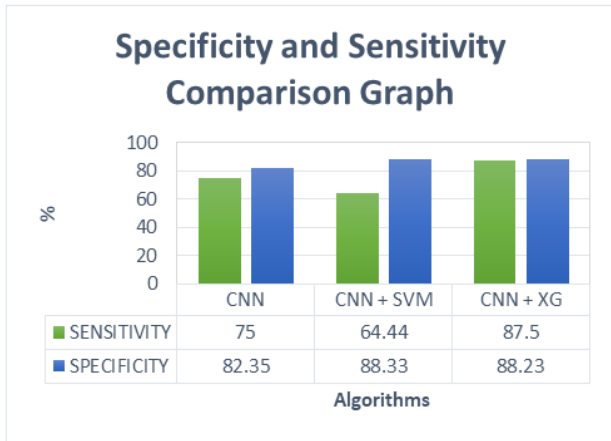
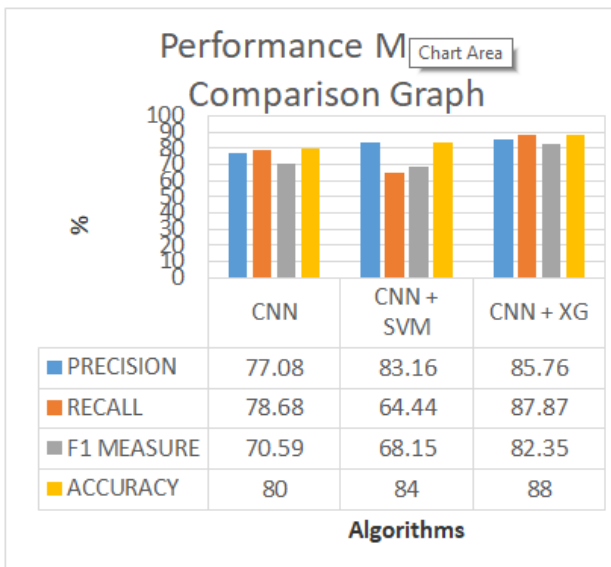*Figure 4. Melanoma sensitivity, specificity and accuracy graph comparison.*

Figure 5. Melanoma sensitivity, specificity and accuracy graph comparison.

Conclusion

The detection of melanoma exactly is incredibly challenging due to low difference among skin and lesion, visual comparability among melanoma and benign, and so forth. Hence, non-invasive technique is required to detect melanoma and benign. To handle these problems, proposed the novel systems to deal with difficulties for exactness and automatic melanoma recognition in dermoscopy images. The framework for melanoma recognition from benign by dermoscopic images using CNN+SVM and CNN+XGB is proposed. The CNN works well for training dataset and SVM is utilized for classification as it is advantageous for generalization. The significant favorable position of SVM is that it gives unified framework where diverse learning machine models can be generated through a proper decision of the portion.

Acknowledgement

Conflict of Interest

The author declares there is no conflict of interest.

References

1. Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. Sensors. 2018; 18(2): 556.

2. Rashmi RP, Bellary S."Review: melanoma detection & classification based on thickness using dermascopic images". IJCTA. 2017;10:821-825.

3. Barata C, Ruela M, Francisco M, et al .Two systems for the detection of melanomas in dermoscopy images using texture and color features. IEEE Systems Journal. 2014; 8(3) :965-979.

4. Yu L, Chen H, Dou Q, et al. Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging.2017; 36(4) : 994-1004.

5. Adegun AA, Viriri S."Deep learning-based system for automatic melanoma detection". IEEE Access. 2020; 8: 7160-7172.

6. Abuzaghleh O, Barkana BD, Faezipour M. Noninvasive real-time automated skin lesion analysis system for melanoma early detection and prevention. IEEE Journal of Translational Engineering in Health and Medicine. 2015; 3: 1-12.

7. Yang J, Xie F, Fan H, et al. Classification for dermoscopy images using convolutional neural networks based on region average pooling. IEEE Access.2018; 23 : 65130-65138.

8. Mane DT, Kulkarni UV. Visualizing and understanding customized convolutional neural network for recognition of handwritten marathi numerals. Procedia Computer Science. 2018; 132: 1123-1137.

9. Patil R, Bellary S. Machine learning approach in melanoma cancer stage detection. J. King Saud Univ - Comput. Inf. 2020; 1319-1578.

10. Hagerty JR, Stanley RJ, Almubarak HA, et al. Deep learning and handcrafted method fusion: higher diagnostic accuracy for melanoma dermoscopy images. IEEE journal of biomedical and health informatics. 2019 ;23 :1385-1391.

11. Mr. S. M. Sangve , Ms. Patil Rashmi R. Competitive analysis for the detection of melanomas in dermoscopy images, nt. J Eng Res Technol.2014; 3 :351- 354.

12. Winkler JK, Sies K, Fink C, et al. Melanoma recognition by a deep learning convolutional neural network-performance in different melanoma subtypes and localisations. Eur J Cancer. 2020;127:21-29.

13. Nida N, Irtaza A, Javed A, et al. Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy c-means clustering. Int. J. Med. Inform.2019 ;124:37-48.

14. K. Korotkov, Quintana J, Campos R, et al. An improved skin lesion matching scheme in total body photography. ieee journal of biomedical and health informatics.2019; 23 : 586-598.

15. Pehamberger H, Steiner A, Wolff K. In vivo epiluminescence microscopy of pigmented skin lesions. i. pattern analysis of pigmented skin lesions. J Am Acad Dermatol. 1987;17: 571-583.

16. Chim H, Deng X. Efficient phrase-based document similarity for clustering. ieee transactions on knowledge and data engineering. IEEE Transactions on knowledge and data engineering. 2008; 20(9): 1217-1229.

17. Barata C, Ruela M, Francisco M, et al. Two systems for the detection of melanomas in dermoscopy images using texture and color features. IEEE Systems Journal. 2014;8:965-979.

18. Ma Z, Tavares JMRS. A novel approach to segment skin lesions in dermoscopic images based on a deformable model. IEEE Journal of Biomedical and Health Informatics. 2016;20:615-623.

*Correspondence to

Rashmi Patil

Department of Computer Science and Engineering

Rao Bahadur Y Mahabaleswarappa Engineering College

Ballari

India

E-mail: rashmiashtagi@gmail.com