Binary Classification of Melanoma Skin Cancer using SVM and CNN

Riya Tanna
Department of Electronics and Communication Engineering
Nirma University
Ahmedabad, India
tanna.riya03@gmail.com

Toshita Sharma
Department of Electronics and Communication Engineering
Nirma University
Ahmedabad, India
toshita2000@gmail.com

2021 International Conference on Artificial Intelligence and Machine Vision (AIMV) | 978-1-6654-4211-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/AIMV53313.2021.9670894

***Abstract*—Skin cancer is seen as one of the most hazardous form of cancers found in humans. Malignant Melanoma is a deadly and a dangerous type of skin cancer. Most skin cancers either spread to other parts of the body and are fatal unless identified and treated early. Medical technology has shown advancement in computer aided diagnosis systems which can classify dermoscopic images. In this paper, we propose two methods for the detection of Skin Cancers particularly with image data taken for melanoma cancerous cells. One is using Convolutional Neural Networks with three layers and the second one is simple model of Support Vector Machines with the default RBF kernel. After applying the image processing techniques, the extracted feature parameters are used to classify the image as Benign or Malignant. The calculation metrics are accuracy, ROC curve and the AUC and confusion matrix. The classification accuracy obtained using SVM classifier is 79.39% and AUC is 0.81. CNN is computed for 100 epochs and the accuracy obtained is 84.39%. The CNN model is bought to deployment in form of a web app with the help of Streamlit.**

Keywords— Melanoma, CNN, SVM, Benign, Malignant, Classification, Skin cancer

I. INTRODUCTION

Among all the types of skin diseases, skin cancer is found to be the deadliest kind of disease in which there is an abnormal growth of melanocytic cells in the skin. Skin cancers are usually caused to people who have exposure to UV light radiations. The malignant tumours are formed due to spreading of skin cells rapidly. Early detection and proper treatment become crucial to prevent it from spreading to the other parts of the body and minimise the death rates. The primary step in the diagnosis of a malignant lesion or sample by a dermatologist is visual examination and without additional technical support, the accuracy rate is around 65%-80%. For complicated cases, the examination is done by taking dermatoscopic images using a high resolution-magnifying camera and some additional screening tests, biopsy tests can further increase the accuracy of the diagnosis. But these processes are expensive and time consuming. Therefore, the focus moves to automated machine learning and deep learning techniques which can classify the cancerous cells as benign or malignant by analysing the results for the images.

Skin cancer is primarily categorised into three types which are as Basal cell carcinoma (BCC), Melanoma and Squamous cell carcinoma (SCC) [Poornima & Shailaja, 2017]. BCC grows slowly and rarely spreads to other parts of the body while SCC is more likely to spread than BCC. Melanoma caused by melanocytes produce a pigment melanin which gives skin its color and it is the most serious type of skin cancer. This paper proposes methods to detect skin cancer and analyse the risk of Melanoma using dermatological images taken with a standard camera. The system will increase the accuracy and reduce the diagnosing time based on machine learning and deep learning algorithms. Convolutional Neural Networks (CNNs) generally seem to automate the diagnosis process producing equal or more accuracies as compared to the conventional methods. CNN is generally used to extract higher level representation of the image content.

The organization of the paper in the subsequent sections will follow the order given: section II presents the review of the related work. In section III, skin cancer classification is done using two machine learning algorithms which are Support Vector Machines (SVM) and Convolutional Neural Networks (CNN). In section IV, experimental results are portrayed on compiling of the modes mentioned above and section V draws the conclusion and discusses the future scope.

II. BACKGROUND AND RELATED WORK

[Poornima & Shailaja, 2017] followed the approach of first segmenting skin cancer images, followed by feature extraction using LBP algorithm and classification is done using SVM classifier based on the features extracted.
In [Daghir et al., 2020], the authors perform pre-processing steps for removing artifacts like hair pixels present in the images. Color enhancement techniques are applied on the images and detection of hair lines is done based on the 2D derivatives of Gaussian distribution. Otsu thresholding method is performed to remove the hair lines from the background of an image. Lesion borders are determined using lesion segmentation process by thresholding and further hair removal methods are incorporated. For model training, 3 different classifiers are used such as SVM, KNN and CNN. KNN classifier has the lowest accuracy which is 57.3%, while SVM performs better than KNN with accuracy of 71.8%. CNN is the most powerful classifier out of the three with the highest accuracy of 85.5%.
[Victor & Ghalib, 2017] applies preprocessing methods which uses median filter and histogram equalizer for image enhancement. Hair removal is done with morphological operators and segmentation is performed using Otsu thresholding. Pre-processing parameters like PSNR, SNR and MSE are also calculated. Accuracy was determined for four classifiers, namely SVM (93.7%) followed by KNN (92.7%), Decision Trees (89.5%) and Boosted Trees (84.3%).
In [Vijayalaxmi ,2019], the pre-processing steps performed are similar to [Daghir et al., 2020 & Victor & Ghalib, 2017] and the model is trained using Back propagation algorithm, SVM and CNN. The algorithm applied in [Vidya & Karki, 2020] performs feature extraction using ABCDE rule which is based on 5 characteristics, namely asymmetry(A),

border(B), color(C), diameter(D) and evolution(E). Image segmentation is implemented using Geodesic Active Contour (GAC) and HOG-GLCM is used for to carry out textural features. Machine learning algorithms such as KNN, SVM and Naïve Bayes classifier are applied and the best accuracy of 97.8% was obtained for SVM classifier with AUC of 0.94. The values of sensitivity and specificity using KNN are 86.2% and 85% respectively.

In [Jusman et al., 2021], multi-layer perceptron, CNN and transfer learning is used to perform skin cancer classification. The HAM1000 dataset is used to train the neural network and the model performs multiclass classification of skin cancer cells. VGG-16 model is used in pre-trained model (transfer learning) while the CNN model consists of 4 convolutional layers. The highest accuracy was obtained by VGG-16 transfer learning model followed by CNN model and then MLP model.

The model in [Lin & Lee, 2020] uses BCN_20000 and HAM10000 dataset to classify 8 categories of skin lesions. Image preprocessing includes balance adjustment, resizing, pixel data normalization and data augmentation for training. The second stage involves CNN models such as ResNet50, MobileNet, DenseNet121, EfficientNetB2, SE-inceptionV3 and VGG-16. CNNs ensemble analysis using Meta Classifier is done in the final stage. The maximum accuracy obtained was 91%.

III. Methodology

This section describes the collection and pre-processing of the dataset followed by the modelling using the algorithms namely SVM and CNN. The data has been divided into training and testing directories which is read into the jupyter notebook using the os.path.join() function in the os library.

A. *Description of Dataset*

The ISIC (International Skin Imaging Collaboration) dataset contains images of benign skin moles and malignant skin moles. The dataset is available on ISIC Archives and Kaggle. It contains colored images each having dimensions 224 x 224. The model is trained using 1440 benign mole images and 1197 malignant mole images while the testing data has 360 benign skin mole samples and 300 malignant skin mole samples.

B. *Data Pre-processing*

The first step is to create labels by using NumPy arrays for the training images. The next part is normalization which is a pre-processing step meaning standardization. It basically means to have different sources of data in the same range. This is done when we divide the array pixels by 255. This brings it in the range of [0,1].

Fig. 1. shows the flowchart for the overall methodology followed from the pre-processing to the final model prediction, providing floating insights on the models used and the classifier.

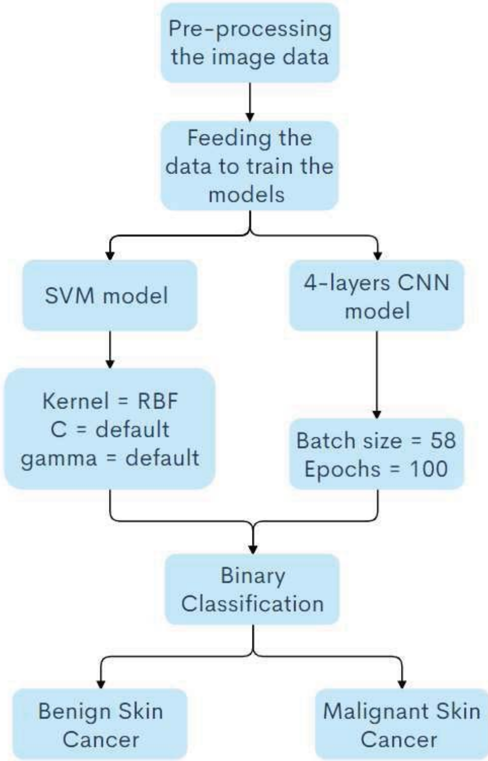

Fig. 1. Flowchart for model implementation

C. *Support Vector Machines*

SVM is a supervised machine learning model which is generally used on a dataset to classify it into varied groups. The algorithm uses a hyperplane which acts like a decision boundary and divides the data into classes. Here, SVM is used to classify the skin cancer images into benign and malignant. SVM is used for optimization of a classification cost and it provides a unified framework in which different machine learning models can be produced through an appropriate selection of kernel [1].

The model is scaled and is trained using a simple SVM classifier. RBF kernel is most widely used kernel and it is generally used for non-linear dataset. It is similar to Gaussian distribution. The RBF kernel function for two points X_1 and X_2 can be mathematically represented as:

$$K(X_1, X_2) = exp\left(-\frac{\|(X_1 - X_2)\|^2}{2\sigma^2}\right) \quad (1)$$

$\|X_1-X_2\|$ is the Euclidean distance between the two points and σ is the variance. The value of RBF function ranges from 0 to 1. σ is also an important parameter to decide the similarity between the two points. The RBF kernel decreases exponentially on increasing the distance between the two points. The width of the region of similarity changes as σ changes. The curve has narrow width for low values of σ but on increasing the value of σ, the points which are farther away are also included in the region of similarity. RBF kernel overcomes the space complexity problem observed in KNN algorithm by storing the support vectors to train the model.

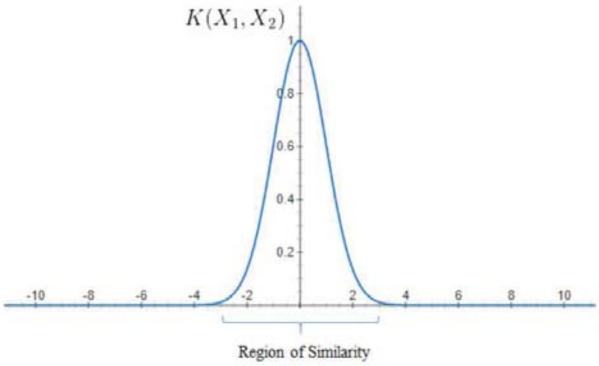

Fig. 2. RBF Kernel for SVM

The hyperparameters of SVM algorithm are C and γ (inversely proportional to σ) which are adjusted according to the dataset to achieve best bias-variance trade-off. The metrics used for evaluation are the confusion matrix and an accuracy score. The accuracy obtained is about 79.4% in the model.

D. Convolutional Neural Networks

Convolutional Neural Networks takes raw pixel data of an image as input and learns how to extract its features unlike machine learning algorithms which requires to pre-process the data to derive features like textures and shapes. For a dataset with images, a kernel/filter of n x n matrix (where n is less than the input image dimension) is selected and it is convolved with the input image to produce the convolved feature matrix of the first layer. The kernel matrix performs element-wise multiplication with input matrix and results to a convolved feature matrix. For a colored image, this step will be performed for R, G and B layers separately and their individual outputs for a particular (i,j) will be added to get a resultant convolved feature matrix. On increasing the number of filters applied to the input, CNN can extract a greater number of features but the training time also increases. So, optimum number of kernels are chosen to take out necessary features for image classification.

The convolutional neural network is a 4-layer structure that takes image height and image width as input parameters and their value is set to 112. This model has been trained with training images and the testing done using validation directory images. With the batch size being 58 and the epochs set to a 100 count, the data is first augmented. Data is augmented to generate new image data with rotations, shearing, changes in scale, translations and horizontal and vertical flips which are geometrical transformations while not changing the original label, used for increasing the accuracy of the model. The ImageDataGenerator is used for performing this function. In the next step, the CNN model applies ReLU activation function to transform the convolved feature. ReLU function has capability to learn faster and it is popular among binary classification. The ReLU function can be mathematically represented as:

$$f(x) = \begin{cases} x \,; x > 0 \\ 0 \,; x \leq 0 \end{cases} \tag{2}$$

The padding has been set to "same" for all three layers to ensure the output dimensions matches the input. The model is compiled with binary cross entropy loss function and the Adam optimizer. The loss function is given by:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log\big(p(y_i)\big) + (1 - y_i) \cdot \log\big(1 - p(y_i)\big) \tag{3}$$

Where y is the label (1/0) which shows binary classification and $p(y_i)$ is the predicted probability. After the ReLU function comes the pooling layer which downsamples the convolved feature matrix while keeping the important feature information. The most common pooling algorithm is max pooling where a size of N x N pooling filter is selected which slides over the input feature map and extracts the maximum value by discarding other values resulting to output feature map.

The class mode is taken as binary because it differentiates between malignant and benign images. The model is using the layers – input layer, hidden layer and an output layer with ReLu function. Convolution is implemented which extracts the only information we want from the image and pooling is used to enhance the power of convolution. The optimizer used is Adam which provides an adaptive learning rate. Graphs are plotted for the training and validation accuracy and loss obtained over the epochs. The accuracy gained is about 84.4%. The model predicts whether the skin mole is benign or malignant based on probability. The model is saved using keras.savemode() and it is deployed using Streamlit, an efficient web app creating library on local host.

IV. Experimental Analysis and Results

The first model trained was using a simple SVM classifier. The results are analyzed using a confusion matrix using a confusion matrix which is as shown below:

True Negative: 227

False Negative: 3

True Positive: 297

False Positive: 133

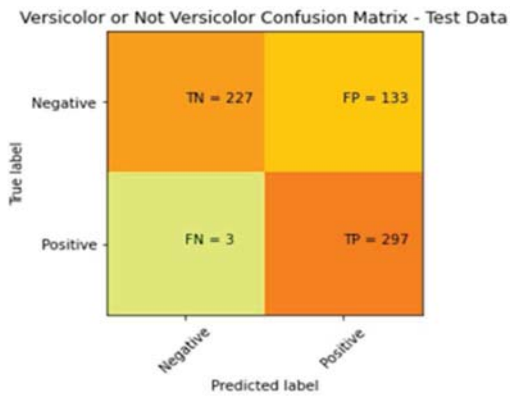

Fig. 3. Confusion Matrix (SVM)

$$Precision = \frac{TP}{TP + FP} = 0.69 \tag{4}$$

$$Recall = \frac{TP}{TP + FN} = 0.99 \tag{5}$$

ROC is receiver operating curve which is one of the model measuring metrics. The area under curve is called the AUC. The higher the AUC, the better is the performance of the model. It works on true positive and negative values which means it shows the ability to perform classification. Here the AUC is 0.81 which means that it can distinguish positive class values from negative class values efficiently.

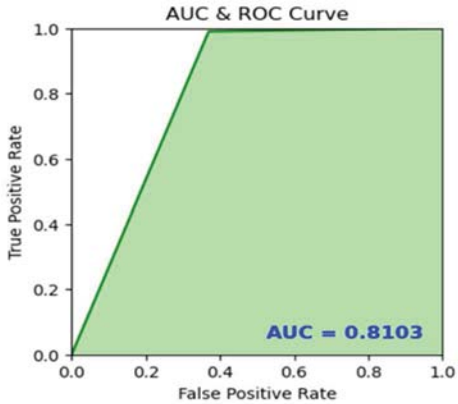

Fig. 4. ROC curve with AUC value

The graphs show training and validation accuracy and loss over 100 epochs.

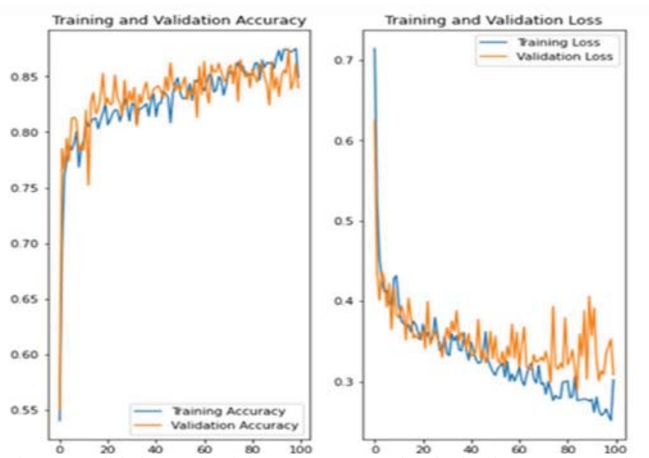

Fig. 5. Validation and training accuracy graphs obtained over 100 epochs

For testing any new image, it has been first converted into a pixel array and then normalised. The entire model has been deployed on local host using Streamlit. It is an open-source app framework in python which creates web-apps. It has compatibility with almost all of major libraries in python.

V. Conclusion and Future Scope

The proposed work shows improvement in classification of skin cancer as benign or malignant. The SVM and the CNN model trained obtain accuracies of 79.39% and 84.39% respectively. The accuracies obtained are not sufficiently high and training a model with large dataset requires high computational time. The model shows the accurate results in less time than traditional methods and it can provide assistance to non-dermatologist physicians. The accuracy obtained from the results will be increased post the augmentation process. Additional hidden layers can be added to increase the accuracy. The future work will focus on improving the results based on different features of the images available in the dataset. To reduce the computational time, ResNet50, VGG-16, MobileNet models can be adapted. Ensemble approaches can be implemented to further analyze the models. Generally, medical images based dataset classification is expected to be highly accurate, the dataset can be improved for future work and some medical insights from experts can help us to have improved performance of the model. The model can be further deployed via websites or android applications to help doctors in their applications of detection.

References

[1] Poornima M, Dr. shailaja K2; Detection of Skin Cancer Using SVM ; International Research Journal of Engineering and Technology (IRJET)

[2] Jinen Daghrir, Lotfi Tlig, Moez Bouchouicha, Mounir Sayadi; Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach; 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)

[3] Akila Victor, Muhammad Rukunuddin Ghalib1; Automatic Detection and Classification of Skin Cancer ; International Journal of Intelligent Engineering and Systems, Vol.10, No.3, 2017

[4] Vijayalakshmi M M; Melanoma Skin Cancer Detection using Image Processing and Machine Learning; International Journal of Trend in Scientific Research and Development (IJTSRD) eISSN: 2456-6470

[5] Vidya, Maya V. Karki; Skin Cancer Detection using Machine Learning Techniques; 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)

[6] Titus Josef Brinker, Achim Hekler, Jochen Sven Utikal, Niels Grabe; Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review; Internet Res. 2018 Oct; 20(10): e11936. Published online 2018 Oct 17. doi: 10.2196/11936

[7] Seetharani Murugaiyan Jaisakthi, Palaniappan Mirunalini and Chandrabose Aravindan, "Automated skin lesion segmentation of dermoscopic images using GrabCut and k-means algorithms", *IET Computer Vision*, vol. 12, no. 8, pp. 1088-1095, 2018.

[8] Shahkar H. Salih and Shereen al-raheym, "Comparison of Skin Lesion Image Between Segmentation Algorithms", *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 18, 2018.

[9] Yuexiang Li and Linlin Shen, "Skin lesion analysis towards melanoma detection using deep learning network", *Sensors*, vol. 18, no. 2, pp. 556, 2018.

[10] Y. Jusman, I. M. Firdiantika, D. A. Dharmawan and K. Purwanto, "Performance of Multi Layer Perceptron and Deep Neural Networks in Skin Cancer Classification," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), 2021, pp. 534-538, doi: 10.1109/LifeTech52111.2021.9391876.

[11] T. -C. Lin and H. -C. Lee, "Skin Cancer Dermoscopy Images Classification with Meta Data via Deep Learning Ensemble," 2020 International Computer Symposium (ICS), 2020, pp. 237-241, doi: 10.1109/ICS51289.2020.00055.