International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-5, January 2020

Design and Implement of Deep Learning Model to Detect the Melanoma

Patange Srujeeth Kumar, Deepak Sukheja, G. Ramesh Chandra

Abstract*: Detecting Skin lesions on the human body is a big task to the doctors in the initial stage because of the low contrast on the body. This skin cancer can be occur due to sun rays. If the disease cannot detect in early stage, there it may cause death to human lives. Here there are some algorithms to predict the melanoma using deep learning techniques. ISIC International Skin Imaging Collaboration Archive set where it provides various images of melanoma and non-melanoma. There are so many challenges to identify the image with melanoma and non-melanoma types of skin cancer. In this paper we applied hair removal algorithm and k-means clustering algorithm where to remove unwanted substances from the original images. To classify the melanoma and non-melanoma skin cancer, this paper proposed prediction process and sequential CNN architecture.

Keywords* : Melanoma, Skin Cancer, Deep Learning, Classification, CNN, Sequential model.

I. INTRODUCTION

In this real world so many different diseases are spreading around us. Due to atmospheric change the virus spreading rapidly around us. Mostly it effects on the skin when the temperature suddenly changes. Melanoma is one kind of cancer where it spreads on the skin where it can spread entire body if we neglects ourselves. As per article https://www.skincancer.org/skin-cancer-information/melanoma/ melanoma is kind of skin cancer where mostly occurs on the skin. Its symptoms are different from normal mole. It size, shape, colour are different from the normal mole. The colours of melanoma are multicolour. Mostly the colours are grey, red, pink, brown, tan, black, blue And shape is in totally unstructured. Size of melanoma rapidly increase and cannot be stopped until we diagnosis. It is the most dangerous in the human life. There different kinds in the melanoma parts. Melanoma can be occur on any part of the skin. And will basically destroys humans life if cannot take precautions from the doctor. This is most rapidly spreading virus on the skin. Till now an estimated cases of 76,380 and estimate deaths of 10,130 in US from [1]. To examine the abnormality in the skin is very tough and may start from any region on the human body. If we neglect this abnormality, it can cause death to human life. Now a days there are so many techniques came into the picture to cure the diseases like radiation therapy, immunotherapy etc. The difference of melanoma and non-melanoma is its size and colour and border region. Size of melanoma is large and has multiple colour on the skin with rough border whereas the non-melanoma will be single colour and comparatively small. At the initial stage doctors cannot find this melanoma because it is impossible to see with naked. Initially it is in small size, that stage even non-melanoma also looks in same size so it's highly impossible with naked eye. To find melanoma at the initial stage is very typical problem for any doctors but with the help of latest emerging technology like Deep Learning using Convolution Neural Network it is also possible to identify melanoma at earlier stage of original melanoma and original non- melanoma as shown in fig[1] and fig[2].

Melanoma

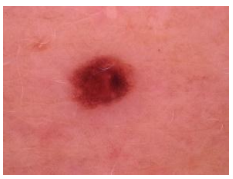

Fig. 1. Non-melanoma

A.Deep Learning:

Deep Learning is a platform of machine learning family. It can be supervised, un-supervised. For the supervised Learning the image classification defines a target set from the classes and train model to recognize with the different images. From the original image it reads the pixels of the image and it find out the difference in the images. CNN model is a technique to recognise different images. Input image been sent to CNN model where it classify and process it.

Manuscript published on January 30, 2020.

* Correspondence Author

Patange Srujeeth Kumar*, Department of Computer Science Engineering, VNR Vignana Jyothi Institute of Engineering & Technology, India. Email: srujeethkumar999@gmail.com

Dr. Deepak Sukheja, Department of Computer Science Engineering, VNR Vignana Jyothi Institute of Engineering & Technology, India. Email: Deepak_s@vnrvjiet.in

Dr. G. Ramesh Chandra, Department of Computer Science Engineering, VNR Vignana Jyothi Institute of Engineering & Technology, India. Email: rameshchandra_g@vnrvjiet.in

System sees that image in array pixel format where it depends on image resolution. With the Image resolution it will check the height*width*dimension. With the help of filters(kernel), pooling and activations CNN model recognise the image.

II. LITERATURE REVIEW

Yading Yuan.*et.al* [1], contrast in images is very low and cannot easily find the affected area and its borders.

So, they proposed fully convolutional neural network with 19 layers architecture on the trained dataset. In this model they used two paths convolutional and deconvolutional. Contextual data combines with the convolutional and maxpooling in convolutional path, from there with upsampling forms in the deconvolutional path to form a full data or image. Novel loss function has been developed which is based on jaccard distance for re-weight sample.

Novel loss function:

$$L(\boldsymbol{w}) = -\frac{1}{N_i \times N_j}\sum_{i,j}[t_{ij}\ln p(\boldsymbol{w}|\boldsymbol{x}_{ij}) + (1-t_{ij})\ln(1-p(\boldsymbol{w}|\boldsymbol{x}_{ij}))]$$

where $t_{ij} \in \{0, 1\}$ is the actual class of $\boldsymbol{x}_{ij}$ with $t_{ij} = 1$ tumor and $t_{ij} = 0$ for background.

Jaccard distance:

$$d_J(M, C) = 1 - J(M, C) = 1 - \frac{|M \cap C|}{|M| + |C| - |M \cap C|}$$

Lequan Yu.*et.al* [2], recognising melanoma is difficult when its contrast is very low and cannot find whether it is melanoma or non-melanoma. To overcome this deep convolutional neural network(CNN) has been used where it can get more features to recognise very accurately.When the network goes deeper they developed residual networks to get interact with the degradation ,overfitting. Later they proposed fully convolutional residual network to get skin lesions accurately. They used Fully Convolutional Residual Network (FCRN) to get the pixel-wise prediction where it values skin lesions in the segmentating task and deep residual network(DRN) from fig[3] which is softmax and support vector machine(SVM) used as a classifiers for the classification method where it can provide more specific features on the skin lesions.

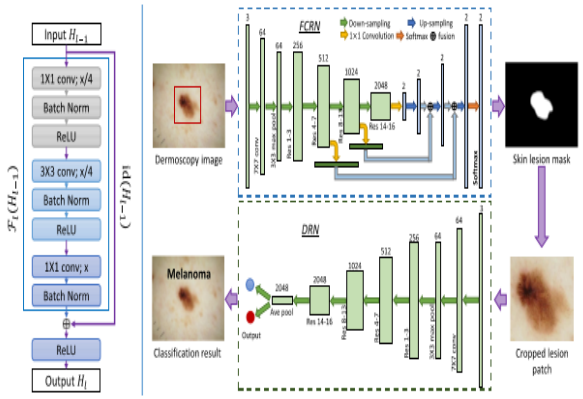

Fig. 2. Architecture of FCRN and DRN for segmentation and classification [2].

Euijoon Ahn.*et.al* [3], defined RSSLS framework, this framework is based on the following components. A. Boundary Connectivity B. Background Detection C. Background Template Creation D. Saliency Measure Via Sparse Reconstruction Error E. Context-Based Error Propagation F. Pixel-Level Sparse Reconstruction Error Creation and finally defined Evaluation Metrics with Dice Similarity Coefficient (DSC).

Lei Bi.*et.al* [4], they proposed two different algorithms image-wise supervised learning(ISL) and multi-scale superpixel based cellular automata(MSCA) to get image contrast high and the background is to be connected well on the image .

Zhen Ma.*et.al* [5], diagnosing skin lesions dermoscopy is a part where it uses imaging technique. There are some skin lesions where its contrast will be low and cannot be identified the skin lesions properly. To handle the skin lesion segmentation they proposed deformable model. The differentiable of normal and skin lesions are seen using RGB color space when it converts.

Jeremy Kawahara.*et.al* [6], on the natural images they proposed linear classifier where it extracts from the trained images using convolutional neural network. Lesions are divided into 10 categories to get the result.Those 10 categories of skin lesions are as following Actinic Keratosis (AK), Basal Cell Carcinoma (BCC), Melanocytic Nevus/Mole (ML), Squamous Cell Carcinoma (SCC), Seborrhoeic Keratosis (SK), Intraepithelial Carcinoma (IEC), Pyogenic Granuloma (PYO), Haemangioma (VSC), Dermatofibroma (DF), and Malignant Melanoma (MEL).

Francesco Peruch.*et.al* [7], proposed with melanocytic lesion segmentation, Mimicking Expert Dermatologists' Segmentations(MEDS) where it provides more accurate result, robustness and speed. Using single parameter tuning MEDS development is simple. Segmentation is very fast in images within fraction of seconds.

M.Emre Celebi.*et.al* [8], lesions are seperated from the background skin using automatic border detection. From this border shapes been extracted. Euclidean distance transform been used to divide images for extracting colour and texture. Using support vector machine the classification been done for the featured data.

III. OBJECTIVES

- To remove unwanted substance from the image and need to be focus on the affected area on the skin using hair removal algorithm and k-means clustering algorithm.
- To diagnosis correctly on the image whether it is melanoma or not using deep learning model.

IV. IMPLEMENTATION

A. Proposed Prediction Process:

Detecting diseases is an important task to the doctors in less time. With naked eye its not so easy to find so there are some algorithms to find the disease and can give accurate results to the doctors. Deep Learning can be used to solve this problem.

Using convolutional neural network and some other pre-processing techniques it can solve the problem. In first part there are few images collected from the International Skin Imaging Collaboration (ISIC). 165 images are collected to the work. So, initially data has been resized and later images are been sent to preprocess where hair removal and k-means clustering takes place.

Hair removal is a technique where it only captures high resolution data in the picture and removes outer shade and only detects infection area on the image. And later these hair removal images are sent to the K-means clustering focus to get colour shades on the image.It focus only on the mostly on the infection area . Using this algorithm the infected area gets brighter than the outerlayer of the infected area which is its surrounding layer. Every image is focused where the infection more occurs and its constrast will be so bright than its surrounded layer of the iamge. In those 165 image dataset, there are 102 images are melanoma and 63 are non-melanoma images. Later these 165 images are seperated for training , testing and validation.

To acheive the above objectives we have proposed prediction process to identify melanoma. The prediction process mentioned in fig[4].

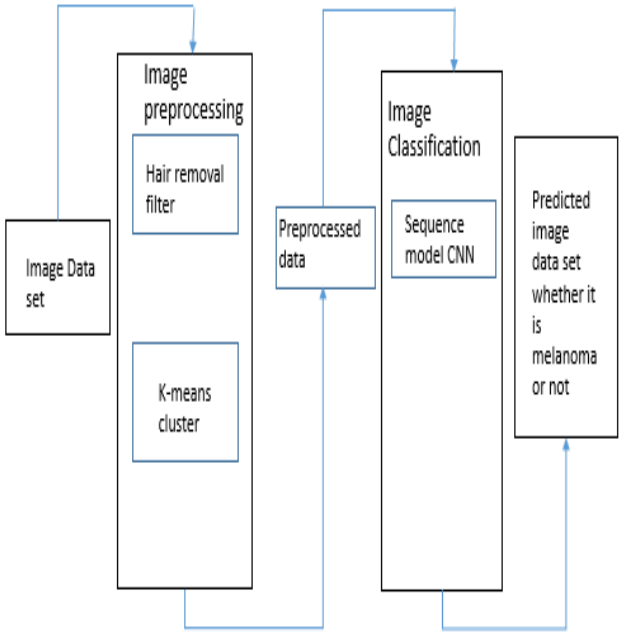

Fig. 3. Melanoma prediction process

B. Implementation:

To implement this proposed prediction process we use windows 10Pro, python 3.7.3, spyder tool as a software and Intel® Core™ i5-6200CPU @2.30GHz @2.40GHz with ram 8GB as a hardware .

Step1: Collected dataset from ISIC archeive. 165 images are collected. In those 165 image dataset, there are 102 melanoma images and 63 non-melanoma images as shown in fig[5].

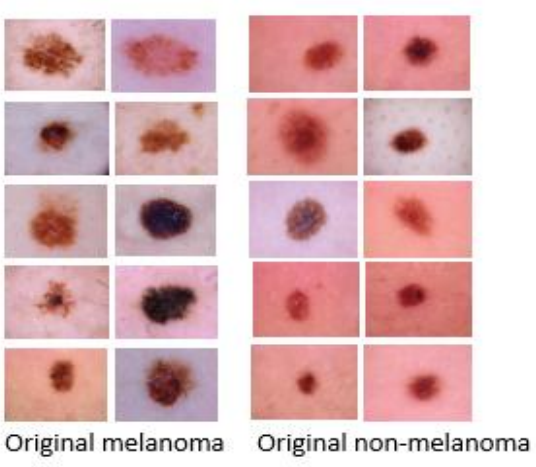

Fig. 4. Original melanoma and non-melanoma

Step2: Images are sent for pre-processing using Hair removal. Using Hair removal algorithm the function only focus on the infection area and removes the unwanted substance from the image as shown in fig[6].

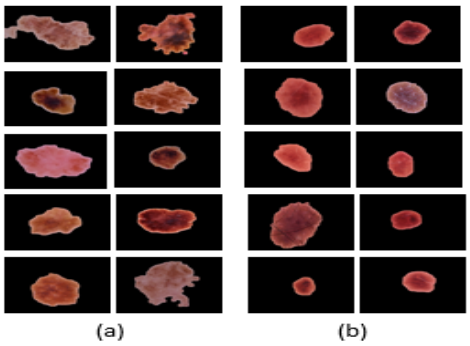

Fig. 5. Hair removal melanoma and non-melanoma in a and b

Step3: Later images are sent for the K-means Clustering. Using k-means clustering algorithm, the images of infected area will get more brighter than the surrounded layer of the image as shown in fig[7].

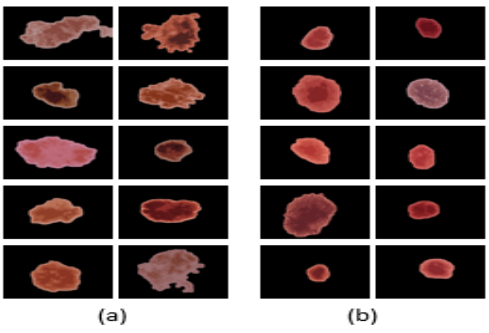

Fig. 6. K-means clustering melanoma and non-melanoma in a and b

Step4: From 165 images 115(71melanoma/44 non-melanoma) images are sent for training.

Step5: From 165 images 50(31 melanoma / 19 non-melanoma)images are sent for testing.

Step6: Using model, few images are classified from the testing whether images are exactly melanoma or not.

C. Dataset:

Collected 165 images from ISIC archeive. 102 images are melanoma, 63 images are non-melanoma from fig[8]. These images are divided for training and testing. For training 70% images(115) are been collected from the total images(165). For training purpose 71 images are melanoma and 44 are non-melanoma images. And for testing 30% images(50) are collected from the total images. In testing 31 images are melanoma and 19 are non-melanoma from fig[9].

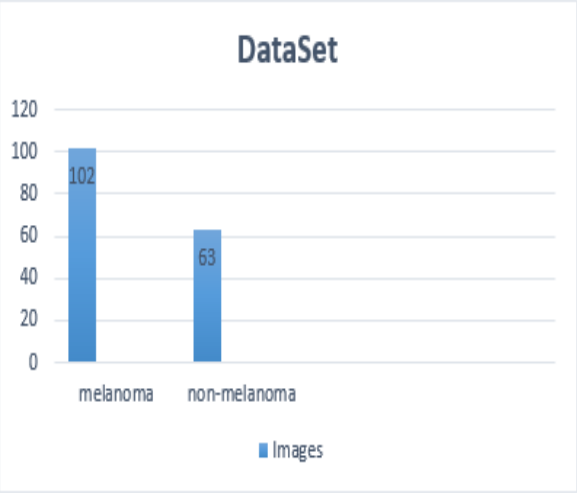

Fig. 7. Image Dataset of melanoma and non-melanoma.

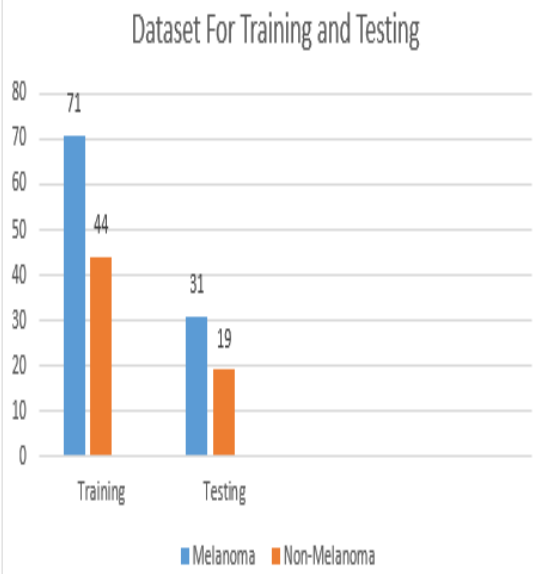

Fig. 8. Traing and Testing dataset.

D. Image Pre-Processing:

During the pre-processing we have used numpy and opencv as a packages. Classifying out a melanoma or non-melanoma from a skin image is a diffcult task. So before going to classifying we pre-processed the images with the some algorithms. They are Hair removal algorithm, K-means clustering algorithm. Hair removal algorithm focused on the affected part in the image and removed unwanted part from the image with background black. K-means clustering algorithm can segment multi colour from the image. Here melanoma can be detect easily when it has multi colour features. To develop this hairremoval and k-means clustering.

a) Hair Removal:

Using this algorithm we will remove the unwanted substances from the image with the help some techniques. The original image is converted to grayscale and later this grayscale image send to threshold where it remove unwanted substances and only focus on affected area. The affected area looks in white colour and unwanted substances looks in black. Later the image send to opening to remove hair around the affected area. And later we send to the dilate where the foreground noise totally removes and object which is affected increases. And mask been created to show how exactly the affected area look in white region with help of pixel array. And finally we send to contours where th object which is shows in colour as seen in orignal image and unwanted substances looks in black as show in fig[10].

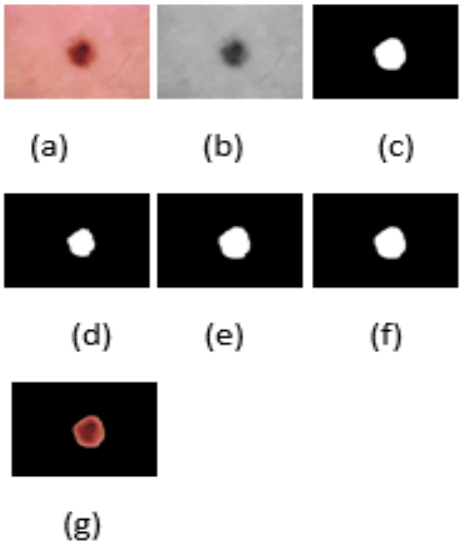

Fig. 9. Hair removal preprocess

Figure 9 shows original image, gray image, threshold image, opening image, dilate image, mask image, final output image with a,b,c,d,e,f and g respectively.

b) K-Means Clustering :

After applying hairremoval to the orginal image. The hair removal image been sent to k-means clustering to know the different colours with help of clusters k.

E. Image Classification:

This process mainly focus on the melanoma and non melanoma from the image data set. K-means clustering is a process where it find the multi colour from the image and can be diagnoised further. Here melanoma can be seen in multi-colour. Here some of the Melanoma and non-melanoma images are sent for training after hair removal and k-means clustering completed. Training data has JPEG file images and along with csv file been created where mentioned as melanoma image as a melanoma and non-melanoma image as a non-melanoma. Some of the training images are also part of the validation randomly. After the training images are trained. New testing dataset images with JPEG file and csv file are sent to the prediction using sequential model 7 layers.

During the implementation of the 7 layers sequential model(CNN Architecture) we have used numpy, pandas, model_from_json to save and load the model, Collections, ImageDataGenerator and keras.layers, sklearn.metrics for accuracy_score.

To classify image whether it is melanoma or not we used deep learning Sequence model. Here sequence model uses the training, validation and testing dataset as a input which are preprocessed images, for the classification to classify whether the image is melanoma or not. This model uses for image classification. Here we are sending images as a input. Here in this model there are 5 convolutional layers and 2 max pooling layer. Input data sent to the model. The 1st layer of convolutional is 32- 3 * 3 filter. In the 2nd layer 32- 3 * 3 filter applied. And then sent to the maxpooling 2*2 size.In the 3rd layer 64- 3 * 3 filter and in 4th layer of convolutional is 64- 3 * 3 and in 5th layer of convolutional is 64-3*3 is applied. And sent to the max pooling layer with 2 * 2 size. And then it sends to flatten layer later it sends to dense 128 units. There are 4 activation filters using relu and 1 activation with softmax, and with the dropout of 0.5 units been used to predict the data whether it is melanoma or not as a output of the model as shown in fig[11].

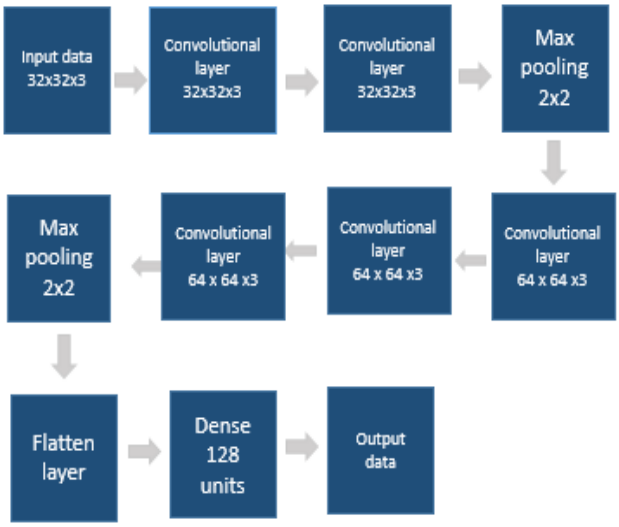

Fig. 10. CNN Architecture

V. RESULT

The workflow of the model of this work shown in fig[12]. To calculate the performance of our proposed model we have conduct experiment in different phases:-

Preprocessed images are sent into three parts training, testing and validation.

Step1 Training:

During this step we used 115 images which are 71 melanoma images and 44 non-melanoma images.

Step2 Testing:

Testing we used 50 images which are 31 melanoma images and 19 non-melanoma images.

Step3 Validation:

Validation is 25% split randomly from the training images.

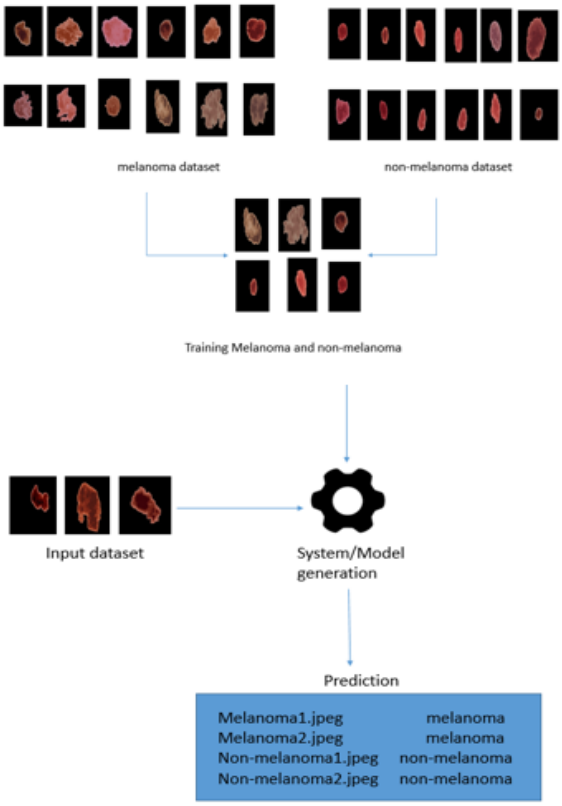

Fig. 11. Workflow of the model

Assumed melanoma as a 0, non-melanoma as a 1 to know how many images are predicting correctly.

MELANOMA TEST DATASET

Before applying algorithm, the original melanoma dataset looks as shown in Table I.

Table I: Before Proposed algorithm on Melanoma

S.no	values
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	0
19	0
20	0

23	0
24	0
25	0
26	0
27	0
28	0
29	0
30	0

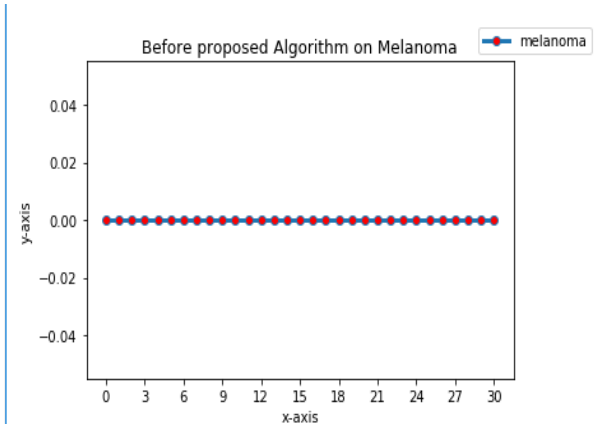

Fig. 12. Before proposed algorithm only on melanoma images

After applying proposed algorithm, the original melanoma dataset looks as shown in Table II.

Table II: After proposed Algorithm on Melanoma

S.no	values
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	1
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	0
19	1
20	0
21	0
22	0
23	0
24	1
25	0
26	0
27	0
28	0
29	0
30	0

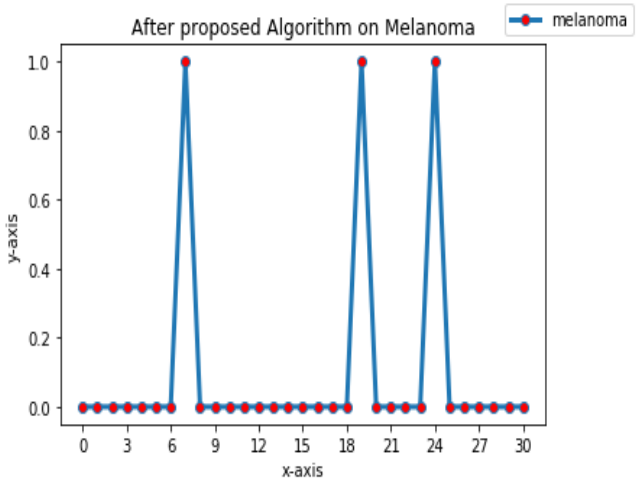

Fig. 13. After proposed algorithm only on melanoma images

NON-MELANOMA TEST DATASET

Before applying algorithm, the original nonmelanoma dataset looks as shown in Table III.

Table III: Before proposed Algorithm on Nonmelanoma

S.no	Values
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1

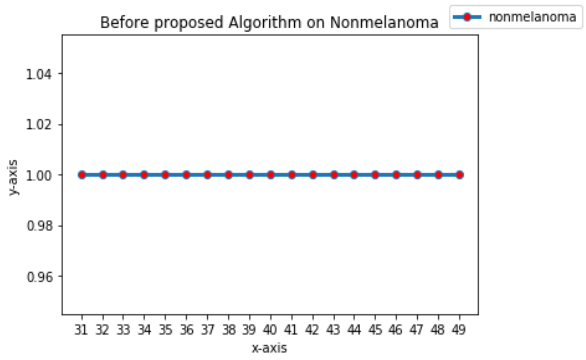

Fig. 14. Before proposed algorithm only on non-melanoma images

After applying proposed algorithm the original no melanoma dataset looks as shown in Table IV.

Table IV : After proposed Algorithm on nonmelanoma

S.no	Values
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1

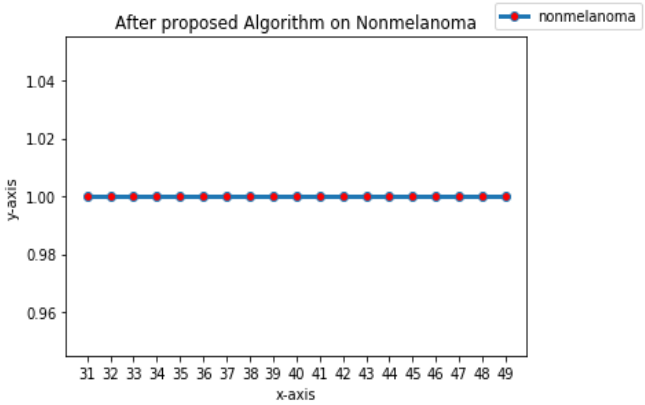

Fig. 15. After proposed algorithm only on non-melanoma images

Accuracy: The accuracy will be calculated on the testing dataset. The mathematical way to calculate accuracy is mentioned below.

$$\text{Accuracy} = \frac{TP + TN}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

Where TP=True Positive, TN=True Negative
FP=False Positive, FN= False Negative

Precision: The precision is a percentage calculation which will be calculated on the testing dataset. Precision is a fractional part of true positives.

$$\text{Precision} = \frac{TP}{\text{TP} + \text{FP}} \quad (2)$$

Recall: The recall is a percentage calculation which will be calculated on the testing dataset. Recall is classified on total number of relevants perfectly correct.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

f1_score: The f1_score will be calculated on the testing dataset.It is a harmonic mean of precision and recall.

$$\text{f1} = 2 * \frac{Precision * Recall}{\text{Precision} + \text{Recall}} \quad (4)$$

Performance of precision, recall and f1_score shown in Table V and fig[17].

Table V: Performance

Support	Precision	Recall	F1_score
31	1.00	0.90	0.95
19	0.86	1.00	0.93

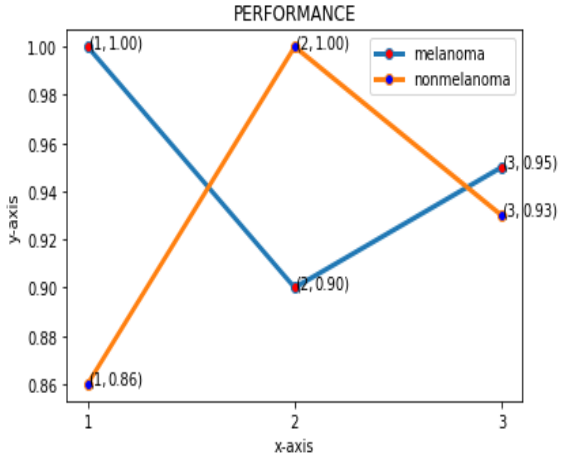

Fig. 16. Performance of Precision, recall, f1

Accuracy Score: 94%

The dataset of the work is 70% on training data set and 30% on testing dataset. Using the sequential model accuracy is 94%.

VI. CONCLUSION

This paper, tends to projected a completely detect melanoma using the deep learning algorithms. Using sequential model the classifcation done on the melanoma and nonmelanoma images after preprocessing technique is applied. The accuracy of this model shows 94%. But there some images are need to be focused more which are not detected, so there may be some other technique for predicting the melanoma more accurately.

REFERENCES

1. Y. Yuan, M. Chao and Y. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," in IEEE Transactions on Medical Imaging, vol. 36, no. 9, pp. 1876-1886, Sept. 2017.
2. L. Yu, H. Chen, Q. Dou, J. Qin and P. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," in IEEE Transactions on Medical Imaging, vol. 36, no. 4, pp. 994-1004, April 2017.
3. E. Ahn *et al*., "Saliency-Based Lesion Segmentation Via Background Detection in Dermoscopic Images," in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 6, pp. 1685-1693, Nov. 2017.
4. L. Bi, J. Kim, E. Ahn, D. Feng and M. Fulham, "Automated skin lesion segmentation via image-wise supervised learning and multi-scale superpixel based cellular automata," *2016* IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, 2016, pp. 1059-1062.
5. Z. Ma and J. M. R. S. Tavares, "A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 2, pp. 615-623, March 2016.).

6. J. Kawahara, A. BenTaieb and G. Hamarneh, "Deep features to classify skin lesions," 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, 2016, pp. 1397-1400.
7. F. Peruch, F. Bogo, M. Bonazza, V. Cappelleri and E. Peserico, "Simpler, Faster, More Accurate Melanocytic Lesion Segmentation Through MEDS," in IEEE Transactions on Biomedical Engineering, vol. 61, no. 2, pp. 557-565, Feb. 2014.
8. M.Emre Celebi,Hitoshi Iyatomi,Gerald Schaefer,William V.Stoecker "Lesion Border Detection in Dermoscopy Images" in ELSEVIER, (2009).

AUTHORS PROFILE

Patange Srujeeth Kumar, currently pursuing Master of Technology in Department of CSE, VNR Vignana Jyothi Institute of Engineering and Technology (VNR VJIET) affiliated to JNTU Hyderabad. Completed Bachelor of Technology in department of Information Technology from JNTU Hyderabad. His work is related to Deep Learning.

Dr.Deepak Sukheja, Associate Professor in Department of CSE, VNR Vignana Jyothi Institute of Engineering and Technology (VNR VJIET). He has 21 Years of teaching and research experience and published more than 23 research papers. His research interest's area includes Data Science , IoT, Data Analytics, Blockchain Technology, Query Optimization, Distributed System, and machine learning.

Dr.G.Ramesh Chandra, Professor in the Department of CSE, VNR Vignana Jyothi Institute of Engineering and Technology (VNR VJIET). He is a Head of Research and Consultancy Centre (RCC) at VNR VJIET. He has 20 Years of teaching and research experience. He published 32 research papers in international journals and conferences. His areas of interest are Image processing, video processing and Data Mining.