SKIN LESION CLASSIFICATION USING DERMOSCOPIC IMAGES

1st D Sai Charan
Computer Engineering
IIITDM Kancheepuram
Chennai, Tamilnadu
coe16b012@iiitdm.ac.in

2nd Subin Sahayam
Computer Engineering
IIITDM Kancheepuram
Chennai, Tamilnadu
coe18d001@iiitdm.ac.in

3rd Dr. J Umarani
Computer Engineering
IIITDM Kancheepuram
Chennai, Tamilnadu
umarani@iiitdm.ac.in

***Abstract*—Skin cancer is the most common cancer in the existing world constituting one third of the cancer cases. A quick and less error-prone solution is needed to diagnose this majorly growing skin cancer. Though there are numerous clinical procedures, the accuracy of diagnosis falls between 49% to 81% and are time consuming. In this project, an automated model for skin lesion classification using dermascopic images has been developed which uses CNN(Convolution Neural Networks). Techniques like data augmentation for tackling class imbalance, segmentation for focusing on the region of interest and 10-fold cross validation to make the model robust have been brought into the picture.**

***Index Terms*—CNN, K-fold cross validation, Data augmentation**

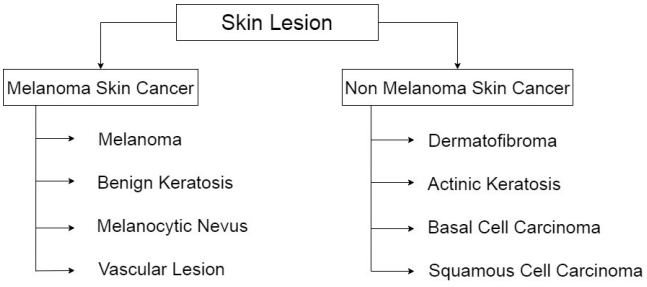
I. Introduction

Skin lesion refers to the worrying and unsettling abnormality on our skin. Skin lesions may be of type cancerous, allergic, etc. Among these, cancer causing skin lesions are hazardous to our health. Some forms of these cancerous lesions are deadly. Melanoma is considered to be the one with a high mortality rate among the cancerous lesions. The occurence rate of melanoma is increasing day by day [1].

Skin cancers are classified majorly into two types. They are

- Melanoma Skin Cancers (MSC),
- Non-Melanoma Skin Cancers (NMSC).

In total, there are 8 types of cancerous skin lesions and are classified as follows :



Clinically, skin cancers are diagnosed based on ABCDE rule [7]. Where,

- A - Asymmetry
- B - Border irregularity
- C - Color of lesion
- D - Diameter of lesion
- E - Enlarging lesion.

This above diagnosis can be done through naked eye. But most skin cancers mimic each other with ABCDE properties.So, there is a possibility of making an error. And there are other clinical procedures like biopsy, etc and are error-prone. On an average, the accuracy of the specialists ranges between 49% to 81%, with one third of melanomas are wrongly predicted as benign lesions [2]–[6]. So dermoscopy came into existence.

Dermoscopy is one of the most important techniques to exam skin lesions and can capture high resolution images of the skin escaping the interruption of surface reflections. Dermoscopy helped in taking more accurate decisions compared to the traditional clinical procedures [8].

Dermoscopy increased the sensitivity, specificity but could not eradicate the probability of making errors and all the above clinical procedures are time consuming, requires human efforts. Early detection of melanoma can save the lives of people. So, a technique is to be created which is fast and less error-prone.

On the other hand, machine learning, especially deep learning based algorithms have become a methodology of choice for analyzing medical images. A deep convolution algorithm can be more objective, accurate and reproducible when it has been well trained. So we came to the conclusion that developing an automated model for skin cancer classification using dermascopic images can aid us in taking quicker, error-less decisions.

This project of developing an automated model,involves creation of different models and usage of different strategies. The dataset is class imbalanced, so data augmentation is done using affine transformations(roate, scale, etc) to generate sufficient numbers of images to tackle class imbalance. Skin lesion images have been segmented to the lesion part, so that the model is made to concentrate more on the lesion region and exclude the surrounding healthy skin. 10-fold cross validation has been brought into the picture. Cross validation makes the model robust. The accuracy will be the mean of the accuracies of all the folds of data.

II. Literature Work

Initially, the major work was done on developing an automated model for diagnosing lesions based on ABCDE properties. In these models [9]–[11], they converted these ABCDE properties of each image into numerical figures, generated a cumulative score by adding all the numerical figures of each property and diagnose the images based on the cumulative score. There is another developed on a seven point checked checklist [12], considered to be an improvisation of ABCDE properties. This research couldn't create the required impact because some skin cancer images mimic the other class images.

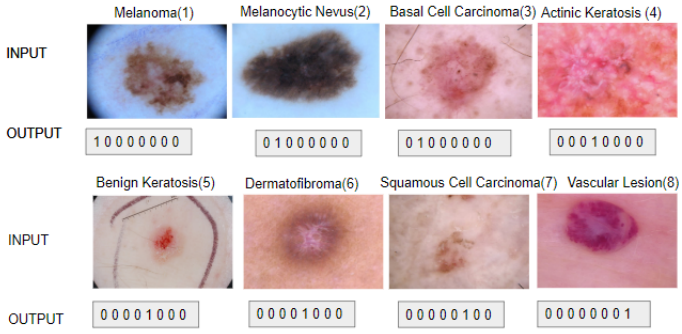
Later, the work was done on classification of skin lesion images using raman spectroscopy. Raman spectroscopy is a noninvasive optical procedure fit for estimating vibrational methods of biomolecules inside practical tissues. In this study, they [13], [14] tested the application of an integrated Raman spectroscopy real time framework for the in vivo diagnosis of skin cancer

In recent times, the major work was carried out on developing models based on CNN variants for classifying skin lesions using dermoscopic images [15]–[17]. They preprocessed the images to eliminate the not interesting regions like eliminating dark circles in the images around the lesion region, contrast balancing, etc. They used data augmentation, random cropping strategies for tackling class imbalance. 5-fold cross validation, Ensembling strategy for making the model robust.

III. Proposed Work

A. Input

The input data set consists of dermoscopic images related to 8 different skin cancers. The expected output of the image is a one hot vector of size '8'.



The dataset is an accumulation of 3 different data sets. i) Ham 10000 (600 x 450 pixels) ii) BCN_20000 (1024x1024 pixels) iii) MSK (various sizes)

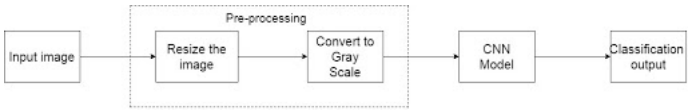
The total count of images is 25331.The count of images for each class are unequal which lead to class imbalance. Each class images' resolutions differ from other class images.

MLN MCN BCC AK BK DF VL SCC
4522 12875 3323 867 2624 239 253 628

Challenges faced:

- Class imbalance
- Varying contrast of images
- Improper images(lesion color is similar to the skin color)
- Different resolutions for images

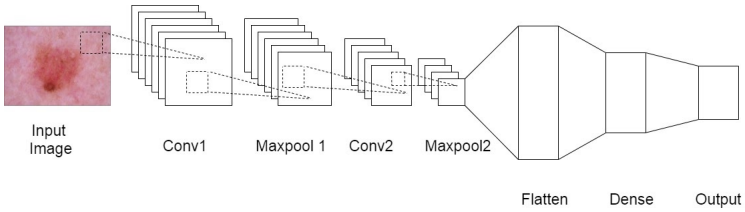
B. Model - 1



This model is developed not to overcome the challenges but to check for the effectiveness of CNN in capturing features of an image. The major concentration is on CNN and it's variants as they are good at capturing neighbourhood properties in an image which is essential in differentiating one skin cancer image from another image.

1) Preprocessing: All the input images are of various sizes so in order to pass these images as input to the model developed they should be of the same size. So, in this phase all the images are scaled down to the size 512 x 512 pixels. These images are further converted into grayscale. This is from our intuition that skin lesion classification does not depend on the color of the lesion.

2) CNN Architecture: This the architecture of the developed CNN Model



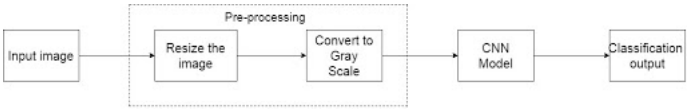
The model is trained on 90 percent of images with optimiser 'Adams' for '20' epochs. The accuracy is 0.42. After going through past work done in the field and clinical procedure for the diagnosis, there are some important issues where model-1 least stressed.

3) Drawbacks:

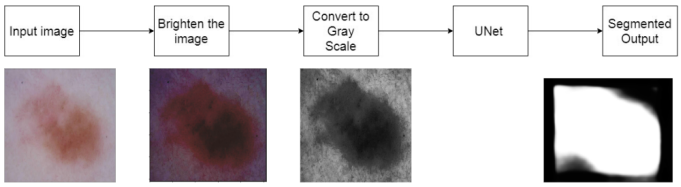
- Not focussing more on regions of interest.
- Color plays a major role in differentiation.
- No tackling for class imbalance.

C. Model-2

This model includes segmentation which helps us in extracting the region of interest and helps us in concentrating more on the region of interest. The input for the model is the same as the input for the model-1.

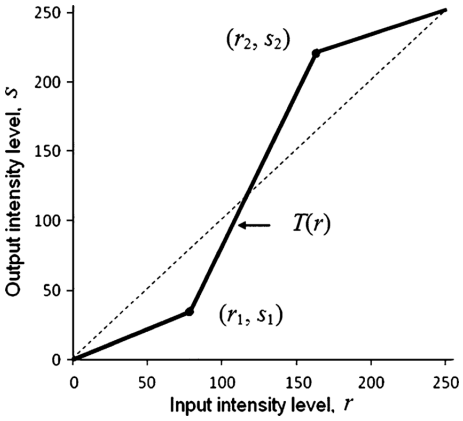


1) Preprocessing: The required lesion region of interest is segmented out from the given dermoscopic images



Steps involved are :

- **Brighten the image :** Color of some class images is so light that there is no distinction between the colors of skin and the lesion. This task is achieved is achieved using 'piece-wise linear transformation' function.



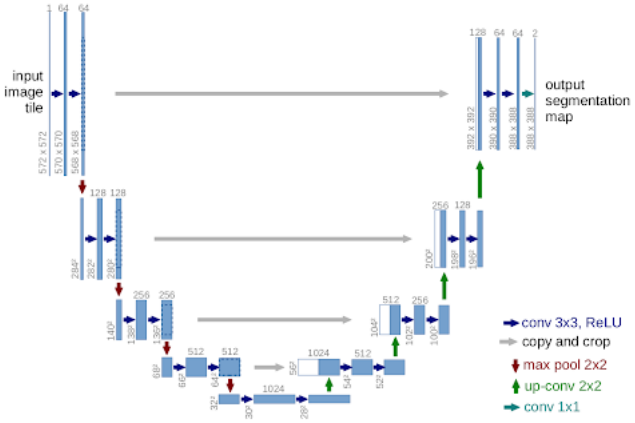
Where,

- (r1,s1) , (r2,s2) as parameters stretches intensity levels,
- Decreases intensity of dark pixels,
- Increases intensity of bright pixels.

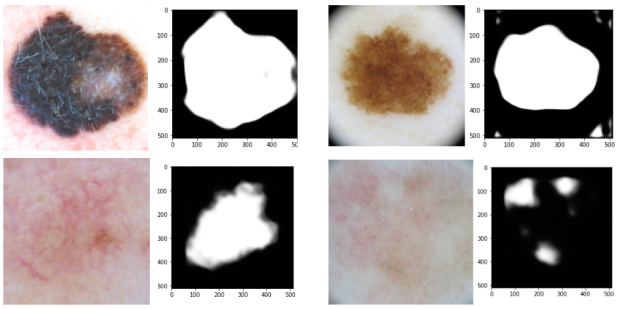
Logic for piece-wise linear transformation :

```
def pixelVal(pix, r1, s1, r2, s2):
if (0 = pix and pix = r1):
return (s1 / r1)*pix
elif (r1  pix and pix = r2):
return ((s2 - s1)/(r2 - r1)) * (pix - r1) + s1
else:
return ((255 - s2)/(255 - r2)) * (pix - r2) + s2
```

- **Grayscale the image :** Convert image to grayscale, this makes the lesion region to stand out with high brightness.
- **U-net Architecture :** U-net,which is a CNN variant is the state-of-the-art architecture for segmentation.

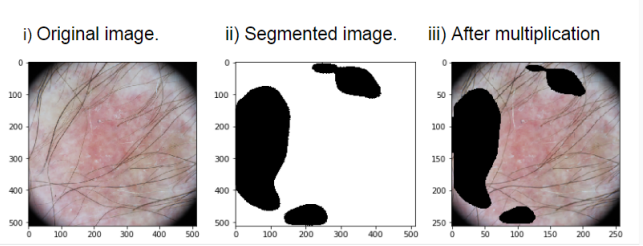


Sample Outputs :

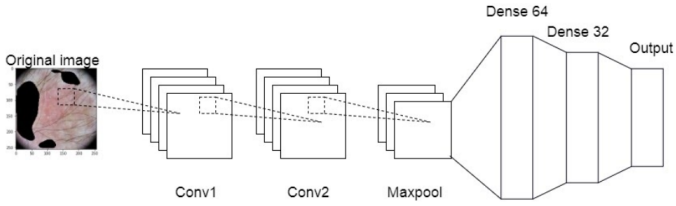


D. CNN Model

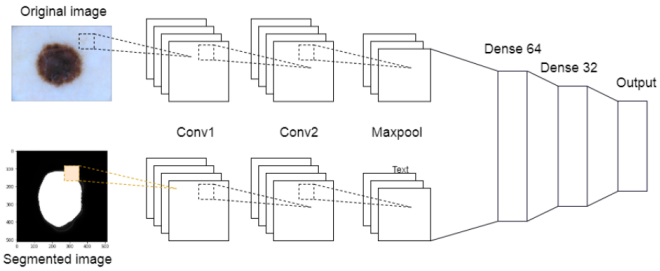
1) Model - 1: This CNN model -1 takes a multiplied image as an input. The multiplied image is created by multiplying the original image with its corresponding segmentation output.



The architecture of Model - 1 that makes use of this multiplied image.



2) Model - 2: This model takes both the original image and segmented images as inputs, there are parallel similar architectures for processing these inputs into the next levels. These images are combined at the flattening stage and are connected to the dense levels.



Output is a one hot vector which points to the predicted class with value '1'.

E. Tackling Data imbalance

The count of images for each class is -

MLN	MCN	BCC	AK	BK	DF	VL	SCC
4522	12875	3323	867	2624	239	253	628

Least size is '239'. So picking 239 images from each class, training the model with dataset of size '239 * 8' images will not be robust. So, additional data is to be created and for that to happen 'Data augmentation' should be brought into the picture. The list of the classes with count lesser than 2000: AK, DF, VL, SCC

So additional data has been created for the above mentioned classes using affine transformations. Affine transformations include rotation, cropping, scaling, flipping(horizontal and vertical), shearing, changing contrast.

F. K-Fold Cross Validation

K-fold cross validation technique makes the model robust. This model is made robust through 10-fold cross validation in which 1-fold is for validation and 9-other folds are for training.

So, the plan is to train the model on the dataset of 16000 images(2000 images from each class). The model is to be trained for 10 times by giving a chance for each fold to act as a validation data set. Find the best model at each time of training and average the accuracies. The mean accuracy will be the accuracy of the model.

IV. Results

Ways of attempt	Model - 1	Model - 2 with CNN model - 1	Model - 2 with CNN model - 2
Without Segmentation	0.204	NA	NA
With Segmentation	0.412	0.561	0.576
With Segmentation, Augmented Data	0.483	0.724	0.732
With Segmentation, Augmented Data, 10 - fold cross validation	0.523	0.863	0.886

Accuracy measure = TP/(TP + FN)

Where,

True positive = predicting the true class correctly,

False negative = predicting the true class incorrectly.

References

[1] Kopf, A. W., Rigel, D. S., Friedman, R. J. (1982). The Rising Incidence and Mortality Rate of Malignant Melanoma. The Journal of Dermatologic Surgery and Oncology, 8(9), 760–761. doi:10.1111/j.1524-4725.1982.tb02676.x

[2] MacKenzie-Wood AR, Milton GW, de Launey JW. Melanoma: accuracy of clinical diagnosis. Aust J Dermatol 1998;39:31–3

[3] Grin CM, Kopf AW, Welkovich B, Bart RS, Levenstein MJ. Accuracy in the clinical diagnosis of malignant melanoma. Arch Dermatol 1990;126:763–6.

[4] Miller M, Ackerman AB. How accurate are dermatologists in the diagnosis of melanoma? Degree of accuracy and implications. Arch Dermatol 1992;128:559–60

[5] Morton CA, Mackie RM. Clinical accuracy of the diagnosis of cutaneous malignant melanoma. Br J Dermatol 1998;138:283–7.

[6] Lindelof B, Hedblad MA. Accuracy in the clinical diagnosis and pattern of malignant melanoma at a dermatological clinic. J Dermatol 1994;21:461–4

[7] Friedman RJ, Rigel DS, Kopf AW: Early detection of malignant melanoma: The role of physician examination and self-examination of the skin. CA Cancer J Clin 35:130-151, 1985

[8] Argenziano, G., Puig, S., Zalaudek, I., Sera, F., Corona, R., Alsina, M., ... Malvehy, J. (2006). Dermoscopy Improves Accuracy of Primary Care Physicians to Triage Lesions Suggestive of Skin Cancer. Journal of Clinical Oncology, 24(12), 1877–1882. doi:10.1200/jco.2005.05.0864

[9] Kasmi, R., Mokrani, K. (2016). Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule. IET Image Processing, 10(6), 448–455. doi:10.1049/iet-ipr.2015.0385

[10] Abbas, Q., Emre Celebi, M., Garcia, I. F., Ahmad, W. (2012). Melanoma recognition framework based on expert definition of ABCD for dermoscopic images. Skin Research and Technology, 19(1), e93–e102. doi:10.1111/j.1600-0846.2012.00614.x

[11] Nilkamal S. Ramteke and Shweta V. Jain. ABCD rule based automatic computer-aided skin cancer detection using MATLAB. Department of Computer Science and Engineering Shri Ramdeobaba College of Engineering Management Nagpur, INDIA

[12] Di Leo, G., Paolillo, A., Sommella, P., Fabbrocini, G. (2010). Automatic Diagnosis of Melanoma: A Software System Based on the 7-Point Check-List. 2010 43rd Hawaii International Conference on System Sciences. doi:10.1109/hicss.2010.76

[13] Lieber, C. A., Majumder, S. K., Ellis, D. L., Billheimer, D. D., Mahadevan-Jansen, A. (2008). In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy. Lasers in Surgery and Medicine, 40(7), 461–467. doi:10.1002/lsm.20653

[14] Lui, H., Zhao, J., McLean, D., Zeng, H. (2012). Real-time Raman Spectroscopy for In Vivo Skin Cancer Diagnosis. Cancer Research, 72(10), 2491–2500. doi:10.1158/0008-5472.can-11-4061

[15] Nils Gessert, Maximilian Nielsen, Mohsin Shaikh, Ren´e Werner, Alexander Schlaefer. Skin Lesion Classification Using Loss Balancing and Ensembles of Multi-Resolution EfficientNets. ISIC 2019.
[16] Steven Zhou , Yixin Zhuang, Rusong Meng. Multi-Category Skin Lesion Diagnosis Using Dermoscopy Images and Deep CNN Ensembles. ISIC 2019.
[17] Federico Pollastri , Juan Maro˜nas , Mario Parreno , Federico Bolelli , Roberto Paredes, Costantino Grana, Alberto Albio. AImageLab-PRHLT at ISIC Challenge 2019. ISIC 2019.