A DEEP LEARNING APPROACH IN SKIN CANCER PREDICTION

[1]Aditi V. Motekar, [2]Dolly A. Soni, [3]Radhika K. Sharma, [4]Saumya R. Agrawal, [5]Prof. P. V. Deshmukh

[1,2,3,4]Students, [5]Assistant Professor
Department of CSE
SSGMCE, Shegaon

***Abstract*: Skin cancer is an alarming disease for mankind. The necessity of early diagnosis of the skin cancer has been increased because of the rapid growth rate of Melanoma skin cancer, its high treatment costs, and death rate. Melanoma, also known as malignant melanoma, is a type of cancer that develops from the pigment-containing cells known as melanocytes. Melanoma accounts for approximately 75% of deaths associated with skin cancer. Melanomas typically occur in the skin but may rarely occur in the mouth, intestines, or eye. In women they most commonly occur on the legs, while in men they are most common on the back. Due to the costs for dermatologists to examine every patient, there arises a need for an automated system to assess a patient's risk of melanoma using images of their skin lesions captured using a standard digital camera.**

In the proposed method the image is processed, segmented and features are extracted. Then the features are compared with the given database and classification is done using artificial neural network. This cancer cells are detected and it takes time to cure in most of the cases. This project proposes an artificial skin cancer detection system using deep learning method. The features of the affected skin cells are extracted after the segmentation of the dermoscopic images using feature extraction technique. A deep learning- based method convolutional neural network classifier is used for the stratification of the extracted features.

I. INTRODUCTION

Skin cancer is the most common human malignancy, is primarily diagnosed visually, beginning with an initial clinical screening and followed potentially by dermoscopic analysis, a biopsy and histopathological examination. Automated classification of skin lesions using images is a challenging task owing to the fine-grained variability in the appearance of skin lesions. Skincancer can be categorized as Melanoma, Squamous cell skin cancer (SCC) and basal cell skin cancer (BCC). From this three-type, melanoma is the most damaging and common type of skin cancer. The main symptom of Melanoma cancer is the appearance of large brown spot, spread by other batches other and darker color, changes in color and size of moles in the body as well as the appearance of dark small and irregular ages, usually hands and feet. As shown in the Fig 1. Malignant melanoma, although far less prevalent than non-melanoma skin cancers, is the major cause of death from skin cancer and is more likely to be reported and accurately diagnosed than non- melanoma skin cancers. Since the early 1970s, malignant melanoma incidence has increased significantly, for example an average 4 per cent every year in the United States. A large number of studies indicate that the risk of malignant melanoma correlates with genetic and personal characteristics, and a person's UV exposure behavior. The purpose of this paper is to explore the efficiency of medical images especially skin cancer images by using the Deep learning neural network. Convolutional Neural Network (CNN) is one of deep learning's methods that has the most significant result in image recognition because it tries to imitate the same way of recognizing images in visual cortex as humans so that they are able to process the same information. The aim of this research is to build a system that can classify melanoma cancer through the images from the dermoscopic examination with Deep Learning training using the CNN method. In the rest of paper, we have shown the literature survey in Section II, in Section III the Analysis of System is presented, in section IV proposed methodology is explained. And in the V section Algorithms and methods are presented. The experiments and results related to data of melanoma skin cancer are also shown in SectionVI. The last section is conclusion and future work of our research. Skin cancer is an alarming disease. So, there arises a need for an automated system to assess a patient's risk of melanoma using images of their skin lesions captured using a standard digital camera.

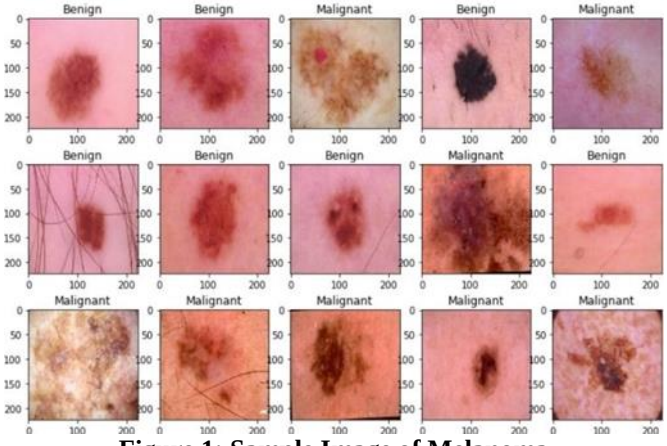

Figure 1: Sample Image of Melanoma

II. Literature Survey

Deep Learning for Skin Cancer Diagnosis with Hierarchical Architectures-IEEE 2019 - In this the results show that a structured classification based on a distinction between malignant and benign lesions, followed by the diagnosis of the latter in different classes leads to better results, when combined with segmented lesions. Skin Cancer Classification using Deep Learning and Transfer Learning- IEEE 2018 - The proposed method outperformed the existing methods of the below listed paper. Skin cancer detection and classification- IEEE 2017 - Convolutional Neural Network (CNN) models can be used for performing this classification without segmentation and extraction of features independently. The input images can be directly fed to the CNN model, which would perform the classification automatically and give better results.

III. Analysis of system

The Proposed system consists of the following architecture

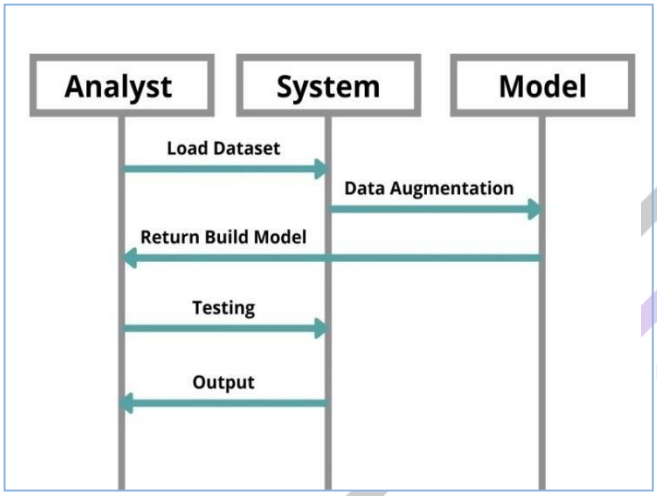

Figure 2: System Architecture

As indicated in the above Fig. 2, initially the dataset will be loaded which has been taken from ISIC; then it will be labeled, shuffled and normalized. To increase the diversity of data available for our training model (without collecting new data), data augmentation strategy will be used which is then followed by the model building in which the data will be processed at different layers of CNN using keras to get the desired output. This process is then followed by cross- validation and testing the proposed model. This step will summarize the performance on a validation/test dataset.The model be classifying the images in backend andthe model will be tested on the accuracy score.

IV. Proposed methodology

A. Dataset

The dataset is taken from the ISIC (International Skin Image Collaboration) Archive. It consists of 1800 pictures of benign moles and 1497 pictures of malignant classified moles. The pictures have all been resized to low resolution (224x224x3) RGB. It has 2 different classes of skin cancer which are listed below:

1. Benign
2. Malignant

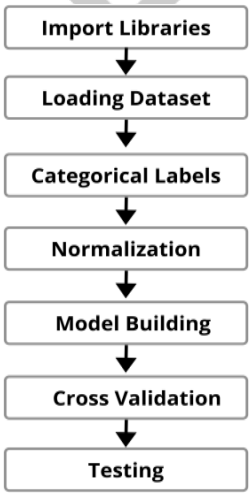

Figure 3: Proposed Methodology

B. Methodology

In this paper our proposed system aids in automatic classification of skin affected image into benign or melanoma by using the CNN classifier. There are some steps to implement the systems which are listed below:

Step 1: Importing the required Libraries

The different libraries are imported. TheOS module in Python provides functions for interacting with the operating system. OS comes under Python's standard utility modules. Numpy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. Scikit-learn is a free machine learning library for Python. It features various algorithms like support vector machine, random forests, and k- neighbor's, and it also supports Python numerical and scientific libraries like Numpy and SciPy. Keras is an Open-Source Neural Network library written in Python that runs on top of Tensorflow. Instead, it uses another library to do it, called the "Backend. Keras is high-level API wrapper for the low-level API, capable of running on top of Tensorflow, CNTK, or Theano.

Step 2: Loading the images from dataset and making Dictionary of images and labels

In this step we load the pictures and turn them into Numpy arrays using their RGB values. As the pictures have already been resized to 224x224, there's no need to resize them. As the pictures do not have any labels, these need to be created. Finally, the pictures are added together to a big training set and shuffled.

Step 3: Categorizing the images and Labeling

Turn the labels in one hot encoding. One hot encoding is a representation of categorical variable as binary vector. The categorical data is the categorical variable whose value takes only values of label. As we know machine learning algorithm works on Binary data that's why we need to categorize them into numerical values.

Step 4: Normalization

Normalize all Values of the pictures by dividing all the RGB values by 255. Normalization is technique often applied as a part of data preparation for machine learning. The goal of normalization is to change the value of numeric column in the data set to a common scale, without distorting differencing in the range of values

Step 5: Model Building

Keras library is used for evaluation of the model the data is posed one by one at different layers CNN to get the desired output. The input layer consist of initial data into system for further processing by the subsequent layer of artificial neurons it is the very beginning of the workflow for the artificial neural network. Convolutional neural layer (conv2Dlayer) it is like set of learnable filter each filter transform the part of image using kernel filter. The kernel filter matrix is applied on whole image. Filter can be seen as the transformation of image.

Step 6: Cross-validation

Cross-validation step consists of creating number of folds. Using the split data function, we will split the dataset in two parts as 70% into training data and 30% testing data for both benign and malignant.

Step 7: Testing

In this step we will test our model by providing the proper input and check that the model is working properly by getting the output as accuracy. The model classifies the images in backend and gives us the output in the form of accuracy. Testing process will end up giving the accuracy of our model.

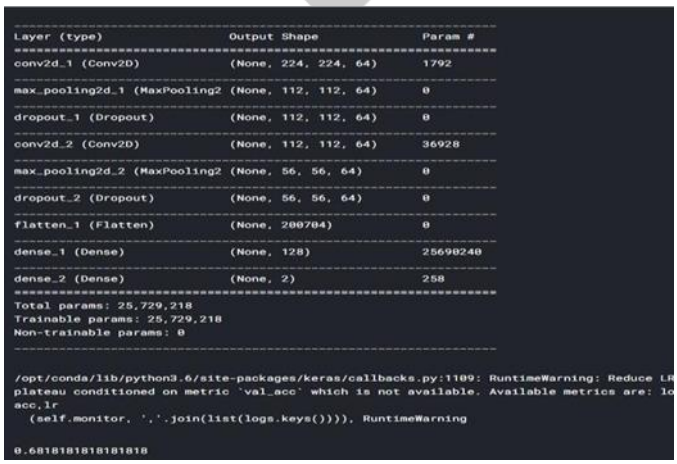

Figure 4: Output

V. Algorithm and Methods

A. Convolutional neural network (CNN)

CNN algorithm is used for categorizing skin image into melanoma or benign. It consists of sequence of several convolutional layers with filters, pooling layers and a fully connected layer. Convolution layer is the initial layer which extracts features from an image. Pooling layer minimizes the size of the activation map. Following are the layers of CNN

1. Feature extraction layer
2. Classification layer
3. Convolutional layer
4. Pooling layer
5. Fully connected layer

CNN is thus a method for transforming the original image layer by layer from the image pixel value into the class scoring value for classification, where each layer has a hyper parameter and some do not have parameters

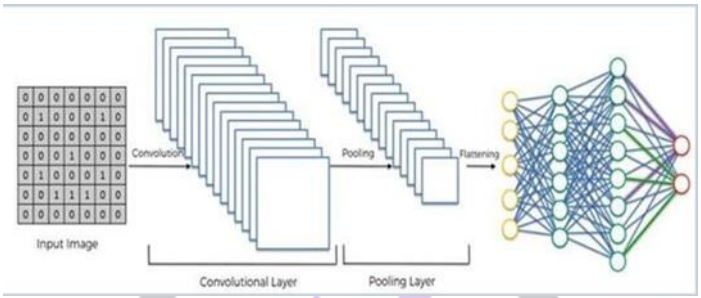

Figure 5: Example of CNN

VI. CONCLUSION

This project proposes an artificial skin cancer detection system using deep learning method. In our proposed system, by combining convolutional and pooling layers, the CNN is able to combine local features and to learn more global features of the image. In this paper, a computational method based on deep learning (Convolutional Neural Network) was implemented, which utilized 3000 images provided by ISIC (International Skin Imaging Collaboration). The proposed method includes images preprocessing for extracting the region of interest in the image itself, and then augmenting some images to produce a bigger dataset which contains 3000 images for each class. The resulting dataset have been applied into CNN model to train the model, which comprise several layers such as convolution layers, pooling layers, and fully connected layers. Testing the model produced promising results with accuracy of 68%. The result encourages and motivate for future improvement and research for online diagnosing of melanoma in early stages before it is too late. Future work of current research is to investigate and restructure the CNN architecture to increase accuracy, obtain more images data for training, and apply new augmentation algorithms to train the model using more data and the ultimate future plan is to make this model accessible and useable using smart phones applications.

REFERENCES

[1] P. Dubal, S. Bhatt, C. Joglekar and S. Patil, "Skin cancer detection and classification," 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI), Langkawi, 2017, pp. 1 -6, doi: 10.1109/ICEEI.2017.8312419.

[2] K. M. Hosny, M. A. Kassem and M. M. Foaud, "Skin Cancer Classification using Deep Learning and Transfer Learning," 2018 9th Cairo International Biomedical Engineering Conference (CIBEC) Cairo, Egypt, 2018, pp. 90-93, doi: 10.1109/CIBEC.2018.8641762.

[3] Agilandeeswari L, Mahajan TanujaSagar, Keerthana N, Skin Lesion Detection using Texture Based Segmentation and Classification by Convolutional Neural Networks (CNN). International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2, December 2019.

[4] V. Pomponiu, H. Nejati, and N.-M. Cheung, ‘‘Deepmole: Deep neural networks for skin mole lesion classification,’’ in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 2623 –2627.

[5] World Health Organization. 2018. Skin Cancers.
Available: http://www.who.int/uv/faq/skincancer/en/index1.ht ml.

[6] L. Deng and Yu D.,―Deep Learning: Methods and Applications‖, Foundations and Trends in Signal Processing, Vol. 7, No. 3 – 4, 2013

[7] Y. LeCun, B. Boser, J.S. Denker, R.E. Howard, W. Habbard, L.D. Jackel and D. Henderson,
―Handwritten digit recognition with a backpropagation network‖, in Advances in neural information processing systems 2, pp. 396-404, Morgan Kaufmann Publishers Inc. San Francisco,
CA, USA, 1990

[8] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016

[9] Y. LeCun, Y., Y. Bengio, dan G. Hinton, 2015.

"Deep Learning", Nature, Vol 521, pp. 436–444, 2015, doi: 10.1038/nature14539
[10] Y. LeCun, L.D. Jackel, Leon Bottou, A. Brunot, C. Cortes, J. S. Denker, H. Drucker, J. Guyon, U. A. Muller, E. Sackinger, P. Simard and V. Vapnik, ― Learning algorithms for classification: a comparison on Handwritten digit recognition‖,1995, http://yann.lecun.com\exdb\ publis\djvu\lecun-95b.djvu
[11] J. Wallen, [Online] Cheat sheet: TensorFlow, an open source software library for machine learning. Available: https://www.techrepublic.com/article/tensorflow- googles-open-source-software-library-for- machine learning-the-smart-persons-guide/., 2017
[12] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M.Blau and S. Thrun, - Dermatologist-level classification of skin cancer with deep neural networks‖, Nature, Vol. 542, pp 115-118, 2017
[13] T.J. Brinker, A. Hekler, J.S. Utikal, N.Grabe, D. Schadendorf, J. Klode, C. Berking, T. Steeb, A.H. Enk, and C.vonKalle, ―Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review‖, Journal of Medical Internet Research, Vol. 20, No.10, pp. 11936- 11946, 201