Prediction of Skin Diseases Using Machine Learning

[1]A. V. Kanade, [2]Mayuri Jadhav, [3]Anushka Jadhav, [4]Mayuri Kulkarni, [5]Durga Nanware

[1]Assistant Professor, Department of Information Technology,Bharati Vidyapeeth's College of Engineering for Women, Savitribai Phule University,India.

[2]UG Student, Department of Information Technology, Bharati Vidyapeeth's College of Engineering for Women, Savitribai Phule University,India.

[3]UG Student, Department of Information Technology,Bharati Vidyapeeth's College of Engineering for Women, Savitribai Phule University,India.

[4]UG Student, Department of Information Technology,Bharati Vidyapeeth's College of Engineering for Women, Savitribai Phule University,India

[5]UG Student, Department of Information Technology,Bharati Vidyapeeth's College of Engineering for Women, Savitribai Phule University,India

Abstract : Recently many deep learning techniques are emerging which enable development of intelligent medical imaging-based diagnosis systems that can assist medical experts in making better decisions. Here we are going to device a model that will focus on skin lesion classification. In this model machine learning techniques will be used for early melanoma detection by classifying the image as malignant or benign. The process of detection will include image processing methods such as image conversion and feature extraction. The processed image will then be used to train a Machine Learning model, which will be developed using Artificial Neural Networks like Back Propagation and Feed Forward networks. This trained model can then be used for detection of an image as cancerous or non-cancerous.

IndexTerms - Natural Language Processing, Machine Learning, Supervised learning, Unsupervised learning , NLP feature, Convolutional Neural Network, FF and BPNN.

I. INTRODUCTION

This Melanoma is one among the deadliest sorts of carcinoma that's caused by abnormal reproduction of melanocyte cells. The incidence of melanoma carcinoma has been increasing over the past few decades. Estimated 76,250 new cases of invasive melanoma were diagnosed in USA in 2012, with an estimated number of 9,180 that end in death. Australia has one among the very best rates of carcinoma within the world. Melanoma is capable of deep invasion. the foremost dangerous characteristic of melanoma is that it can spread widely over the body via the lymphatic vessels and blood vessels. Thus, early diagnosis of melanoma may be a key factor for the prognosis of the disease. However, it's often undiagnosed or misdiagnosed, because it requires biopsy of the skin i.e. a neighborhood of the skin must be extracted and diagnosed so as to detect the cancer. Diagnosis of the disease are often done using dermoscopic images but it requires physicians who have an honest amount of experience. Hence, there's a requirement for a system so as to assist physicians for the classification of melanoma as cancerous or non-cancerous. Here we'll perform detection of cancer using dermoscopic images, which is given as input. We'll then perform image processing after which we'll give the image for training to the Machine Learning algorithm. The Machine Learning model will then be trained on the training dataset. After the model has been trained well, we'll use it for detection of a replacement input image. The new image will again undergo image pre-processing steps. It'll then tend as input to the trained model which will classify whether the input image is that of a malign cancer or that of a benign cancer. Physicians can then use this output to require any longer decisions regarding further diagnosis or analysis.

II. LITERATUR SURVERY

Hiam Alquran Isam, Abu Qasmieh, Ali Mohmmad Alqudah, Sajidah Alhammouri Esraa Alawneh, Ammar Abughazaleh [1] stated that, Skin cancer can be classified into various types such as Melanoma, Basal layer carcinoma, Squamous cell carcinoma among which Melanoma is found to be the most unpredictable. Early detection of Melanoma can be helpful in curing it. Melanoma is capable of deep invasion, so the most dangerous characteristic of it is that it spreads widely over the body. In this, Pre-processing is done by Converting, Contrast enhancement, Histogram modification, Noise Filtering is done.

The second stage after pre-processing is detecting and segmenting the region of interest (ROI) which represents the lesion region. The segmentation stage includes steps: Image threshoulding, image filling, image opening, converting extracted region to gray level, and then performing histogram equalization to the extracted gray level image. After extracting the lesion (ROI) in the segmentation stage, the predefined features will be extracted from the ROI for classification. The selected features are shape, color and various texture features. Since these images have some statistical texture features, it uses one of the common algorithm to extract such features which is Gray Level CoOccurrence Matrix (GLCM). In addition, the dermoscopy features (ABCD) are important in distinguishing skin lesion types. We combined these features to get a good classification results for distinguishing the benign from the malignant skin lesions. The feature extracted are fed into PCA(Principal Component Analysis). The PCA uses the correlation matrix instead of the covariance matrix. After implementing this operation and calculation of eigenvalues and variances, set of main components is obtained which are arranged based on their ability to distinguish between benign and malignant lesions.

SVM is used for the data classification, selected features(TDS, mean, standard deviation, energy, and contrast) using PCA are fed into the SVM model which is used to classify the image into binary classes benign and malignant.

Adria Romero Lopez, Xavier Giro-i-Nieto Universitat Politecnica de Catalunya Barcelona, Catalunya, Spain, Jack Burdick [2] focused on the problem of skin lesion classification, particularly early melanoma detection, and present a deep-learning based approach to solve the problem of classifying a dermoscopic image containing a skin lesion as malignant or benign. The proposed solution is built around the VGGNet (Very Deep Convolutional Network for Large-Scale Visual Recognition)architecture and uses the transfer learning paradigm.
The input image is passed through five convolutional blocks. Each convolutional block includes a 2D convolution layer operation (the number of filters changes between blocks). All hidden layers are equipped with a ReLU (Rectified Linear Unit) as the activation function layer (nonlinearity operation) and include spatial pooling through use of a max-pooling layer. The network is concluded with a classifier block consisting of three Fully Connected (FC) layers.

Input images must be pre-processed by normalizing the pixel values, cropping the image to square aspect ratio, resizing the image to the expected size. To increase the accuracy of the model, the data is augmented via a number of random transformations. The modified VGG16 ConvNet can be used in three different ways: (i) training the ConvNet from scratch; (ii) using the transfer learning paradigm to leverage features from a pre-trained VGG 16 on a larger dataset; and (iii) keeping the transfer learning paradigm and fine-tuning the ConvNets architecture. Keras a deep learning framework for Python, was utilized to implement the neural network architecture

N Vikranth Kumar et al.[3] proposed that Skin infections, for example, Melanoma and Carcinoma are regularly very difficult to identify at a beginning phase and it is considerably harder to group them independently. As of late, it is notable that, the most risky type of skin malignancy among different kinds of skin disease is melanoma since it is significantly more prone to spread to different pieces of the body if not analysed and treated early. To characterize these skin sicknesses, "Support Vector Machine (SVM)" an AI Calculation can be utilized. In this paper, we propose a technique to recognize if a given example is influenced with Melanoma. The means engaged with this examination are gathering named information of pictures that are pre-prepared, smoothing those pictures and getting the pixel powers of pictures into an exhibit, affixing all such clusters into a data set, preparing the SVM with marked information utilizing an appropriate bit, and utilizing the prepared information to order the examples effectively.

Quan Gan et al. [4] stated that the Skin illnesses genuinely affect individuals' life and wellbeing. Ebb and flow research proposes a proficient way to deal with distinguish solitary kind of skin illnesses. It is important to foster programmed techniques to expand the precision of determination for multi type skin infections. In this paper, three sort skin sicknesses like herpes, dermatitis, and psoriasis skin illness could be recognized by another acknowledgment strategy. At first, skin pictures were pre-processed to eliminate clamor and insignificant foundation by sifting and change. At that point the strategy for dark level co-event lattice (GLCM) was acquainted with portion pictures of skin illness. The surface and shading highlights of various skin sickness pictures could be gotten precisely. At long last, by utilizing the help vector machine (SVM) order strategy, three sorts of skin sicknesses were distinguished. The exploratory outcomes show the adequacy and achievability of the proposed technique.

III. PROPOSED SYSTEM

This system basically will classify the dermoscopic images as cancerous or non-cancerous. The input to the system will be cancerous as well as non-cancerous images. After giving the input, RGB to HSV conversion will be performed followed by feature extraction from the images using CNN. After feature extraction, the images will be trained based on features using machine learning algorithm. The machine learning model will be formed from training data set. Training will be followed by detection. RGB to HSV conversion and feature extraction of images will done for the test images. These images will be given to training model for the classification. The problem being a two class classification problem will determine whether a dermoscopic image containing a skin lesion contains a melanoma(malignant lesion) or a benign lesion.

IV. IMPLEMENTATION

For implementing this project, we have divided this problem into following modules:

1. Training Module
   - 1.1 Pre-processing
   - 1.2 RGB to HSV conversion
   - 1.3 Segmentation
   - 1.4 Feature extraction
   - 1.5 FF + BPNN

In the training module, first the image dataset is prepared, then by using CNN image pre-processing is done. The RGB to HSV conversion is done on the image and using CNN's different layers image is filtered and features are extracted. The Extracted features are trained and model is created. Using FF and BPNN algorithm the images are classified.

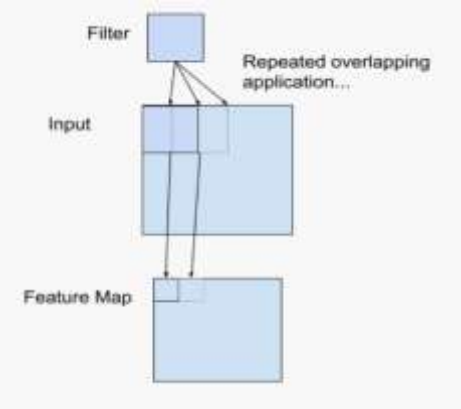

Fig. 1 Example of a filter applied

2. Prediction Module
    1.6 Pre-processing
    1.7 RGB to HSV conversion
    1.8 Segmentation
    1.9 Feature extraction
    1.10 Prediction

In the prediction module, the test dataset is provided as a input to the trained model and using classification algorithm the output isdetected as the image is cancerous or non-cancerous.

System Design

System design is a process of defining architecture, models, interfaces and data for a system to satisfy specified requirements. System design could be seen as the application of system theory to product developing. There is some overlap with disciplines of system analysis, system architecture and system engineering.

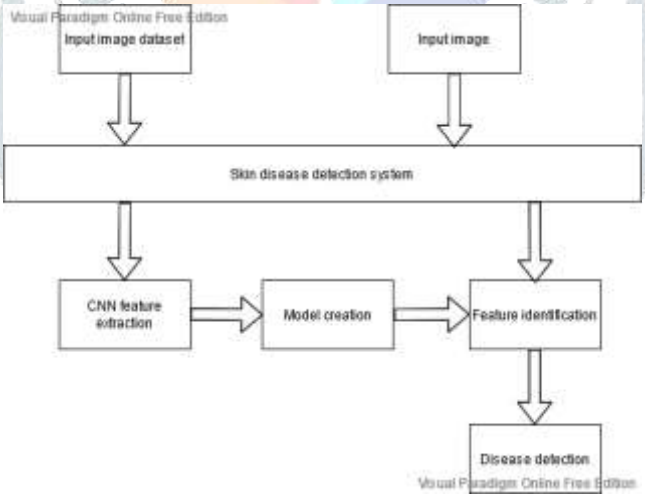

Fig.2 System Architecture

V. RESULTS

Fig. 3 Skin Cancer Image

Input image datasets will be given to the system in the form of dermoscopic image shown in Fig. 3.
Initially, input images are pre-processed by performing operations like RGB to HSV conversion, segmentation, histogram generation. Features are extracted using CNN. And the training model is created.
In the prediction module, the test image is given to the system. The features are extracted from the test image and measured the similarity between the features of the test image and features of the dataset images. And gives the result as the image is cancerious or non-cancerious.
Thus system will be able to detect melanoma cancer of the patient with enough amount of accuracy after being trained with a limited set of data. The system will provide early diagnosis of the disease and will be helpful for such patients as cost will be less as compared to biopsy.

VI. CONCLUSION

In the end we conclude that the application can predict the patients melanoma cancer more accurately than its current state and thus improve its performance. The future work would involve improving its accuracy and may be make it more advanced and user friendly.

REFERENCES

[1] Hiam Alquran1, *, Isam Abu Qasmieh1, Ali Mohammad Alqudah1, Sajidah Alhammouri1, Esraa Alawneh1, Ammar Abughazaleh1, Firas Hasayen1, "The Melanoma Skin Cancer Detection and Classification using Support Vector Machine".

[2] Adria Romero Lopez, Xavier Giro-i-Nieto Universitat Politecnica de Catalunya Barcelona, Catalunya, Spain, Jack Burdick, Oge Marques Florida Atlantic University Boca Raton, FL, USA, "Skin Lesion Classification From Dermoscopic Images Using Deep Learning Techniques".z

[3] N Vikranth KumarVIT University, Vellore, India; P Vijeeth Kumar; K Pramodh; Yepuganti Karuna "Classification of Skin diseases using Image processing and SVM" 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN).

[4] Quan Gan,1 and Tao Ji1 "Skin Disease Recognition Method Based on Image Color and Texture Features" International Research Journal of Engineering and Technology (IRJET) eISSN: 2395-0056 Volume: 06 Issue: 04 | Apr 2019 www.irjet.net p-ISSN: 2395- 0072

[5] Ho Tak Lau School of Electrical, Mechanical and Mechatronic Systems, Faculty of Engineering, University of Technology, Sydney, Australia, Adel Al-Jumaily School of Electrical, Mechanical and Mechatronic Systems, Faculty of Engineering, University of Technology, Sydney, Australia, " Automatically Early Detection of Skin Cancer: Study Based on Neural Network Classification"

[6] Shuying Liu, Weihong Deng Beijing University of Posts and Telecommunications, Beijing, China "Very Deep Convolutional Neural Network Based Image Classification Using Small Training Sample Size"

[7] Codella, N. C. F., Nguyen, Q.-B., Pankanti, S., Gutman, D. A., Helba, B., Halpern, A. C., & Smith, J. R. (2017). "Deep learning ensembles for melanoma recognition in dermoscopy images"

[8] Kawahara, J., BenTaieb, A., & Hamarneh, G. (2016). "Deep features to classify skin lesions". 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI).

[10] Catarina Barata, Margarida Ruela, Mariana Francisco, Teresa Mendonça, and Jorge S. Marques "Two Systems for the Detection of Melanomas in Dermoscopy Images Using Texture and Color Features". This paper aims at determining the best system for skin lesion classification.

[11] Ammara Masood, Adel Al- Jumaily, Khairul Anam " Self-Supervised Learning Model for Skin Cancer Diagnosis"

[12] S. Sabbaghi, M. Aldeen, Senior Member, IEEE, R. Garnavi, Member, IEEE:" A Deep Bag-of-Features Model for the Classification of Melanomas in Dermoscopy Images".