WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Skin Cancer Detection Using Machine Learning

[1]Ms. Vijaya Lakshmi, [2]Sneha S, [3]Ramya R and [4]Haaniya G

[1,2,3,4] Sri Manakula Vinayagar Engineering College, Madagadipet, Pondicherry, India

ABSTRACT

In today's modern world, Skin cancer is the most common cause of death among humans. Skin cancer is an abnormal growth of skin cells most often develops on the body exposed to sunlight, but can occur anywhere on the body. Most skin cancers are curable at early stages. So early and fast detection of skin cancer can save the patient's life with the new technology, and early detection of skin cancer is possible at the initial stage. The formal method for diagnosis of skin cancer detection is the Biopsy method. It is done by removing skin cells and the sample goes to various laboratory testing. It is a painful and time-consuming process. We have proposed a skin cancer detection system using a convolution neural network for the early detection of skin cancer disease. It Is more advantageous to patients. The diagnosing methodology uses an Image processing method algorithm. The dermoscopy image of skin cancer is taken and it goes under various pre-processing techniques for noise removal and image enhancement. Then the image is undergone segmentation. These features are given as the input to the classifier. Convolution neural network is used for classification purposes. It classifies the given image into cancerous or non-cancerous.

Keywords: Malignant melanoma, Keras Sequential API, Biopsy method, Convolution Neural Network (CNN), Optimizer, Learning Rate (LR), Support Vector Machine (SVM)

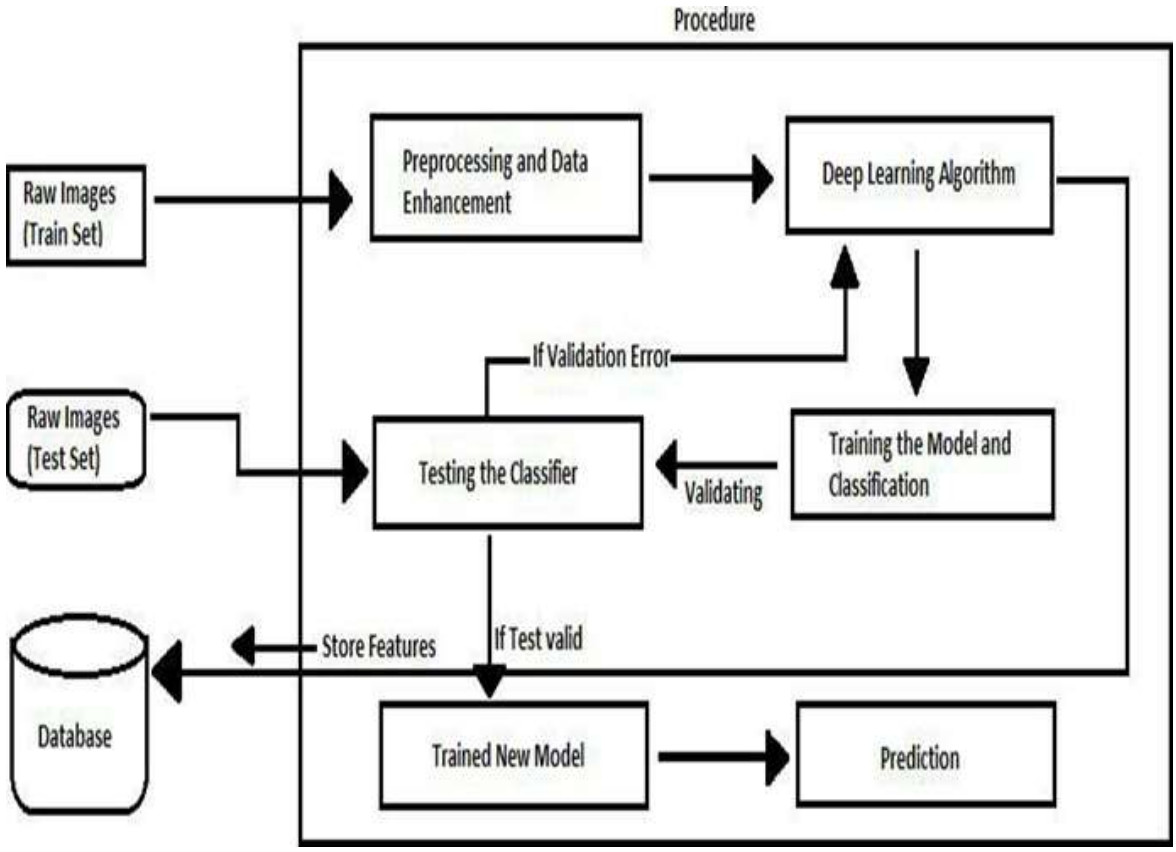
1. Introduction

In Australia, America, and Europe, malignant melanoma has become one of the leading cancers in recent decades. In most cases, the curability of skin cancer is high, reaching over 92% if it is detected early. Often a biopsy is performed to determine whether a tumor is malignant or benign. Since undergoing a biopsy is expensive and causes morbidity, automatic early detection tests are becoming a much faster and more convenient alternative to undergoing a biopsy. Certain physical features and color information typical of different types of skin cancer can be used to automate the diagnosis of skin melanoma, according to dermatology imaging researchers. Vertical thickness, three-dimensional (3D) size and shape, and color of melanoma lesions are the most important diagnostic and prognostic parameters. In addition to irregularities in the lesion's boundary, early melanoma can have non-uniform pigmentation with a wide range of colors. Previously, a machine-learning approach was used to detect skin cancer. A method for detecting skin cancer automatically is presented and discussed in this paper, which uses several approaches to classify cancer images according to whether they are benign or malignant.

The images in the databases used are contained both digital photo and dermoscopy images. The image databases are collected from Sydney Melanoma Diagnostic Centre in Royal Prince Alfred Hospital and the internet website. A total of 448 images are collected and separated into four groups to test the accuracy of each group. The grouping policy is determined by the outlook of the image and its properties. Group A contains 93 digital malignant images from the internet; group B contains 142 dermoscopy malignant images from Sydney Melanoma Diagnostic Centre; group C and D contains 121 and 92 digital benign melanoma images from the internet. Dermoscopy, also calls Dermatoscopy or Epiluminescence light microscopy (ELM). It was first announced in 1987, it is a kind of imaging technique used to exanimate lesions with a dermatoscopy. The process is done by placing an oil immersion between the skin and the optics. The lighting can magnify the skin that improves on reveal most of the pigmented structure, different color shades that are not visible to the naked eye; and allows direct viewing and analysis of the epidermis (the outer layer of the skin) and papillary dermis (the deep vascular inner layer of the skin). A physician uses this technique for the diagnosis of skin cancer more efficiently. Historical research has proved that ELM can improve diagnostic accuracy by 5 – 30% compared to traditional imaging. Furthermore, ELM has evolved and is digitalized that can be used for classification with the computer. It gives several advantages in diagnosing like considering small suspicious lesions and objective evaluation of parameters: geometry, color, and texture; and storage of image and comparison for future development.

2. Proposed System

We have used Keras Sequential API, where you have just to add one layer at a time, starting from the input. Conv2D layer, a set of learnable features. The number of filters used here is thirty-two. Each filter transforms a part of the image which is defined by the kernel size using the kernel filter. Transformed images are the filter maps. The next important layer is the pooling layer which simply acts as a down-sampling filter. Combining both the above layers, CNN gets the ease to combine local features and learn global features. Activation Function relu is used to add non-linearity to the network. We use a



(Block Diagram)

regularization method, where a proportion of nodes in the layer are randomly ignored (setting their weights to zero) for each training sample i.e. the Dropout function. This improves generalizing the network. Now, to convert the final feature maps into one single 1D vector we need to flatten them, thus Flatten Layer is used. This flattening step is needed so that you can make use of fully connected layers after some of the above layers. It combines all the found native options of the previous convolutional layers. In the last layer, Dense () is used which gives the net output distribution of the likelihood of every category. Once layers have been added, we need to set up a score function, a loss function, and a proper optimization algorithm. We define binary cross entropy as our loss function which will measure the error rate between the observed labels and predicted labels. The next most important is the optimizer. Adam Optimizer has an advantage as it involves functions of other optimizers as well. Adam is a well-known and popular algorithm in the field of learning models. Next, is the metric function which is used to evaluate the performance of the system, metric accuracy is used. Learning Rate (LR) is another important term. It is an annealing method. Ideally one should have a decreasing Learning Rate during the rate to have a minimal loss. Reduce LROn Plateau is used; the name itself means to reduce the LR to reach the global minimum of a loss function.

3. Modules description

3.1 Load DataSet:

The dataset is image model training and the test was collected and screened from real diagnostic cases. All images were generated from microscopic observation. Normally, after target tissues are obtained from lesions, the preparation of samples includes the following steps: fixation, selection and trimming, embedding, sectioning, and staining. After a sample is prepared, it is examined under the microscope manually to locate abnormal areas or possible clues for a certain skin disease. All of these images are in the RGB color space, with size. We resized the large images and fixed the image size to 224,224 for model training and testing

3.2 Image Processing:

Achieving high performance of skin disease detection systems requires overcoming some major difficulties. Such as creating a database and unifying image dimensions. In the following section, the technique used in image resizing is explained.

To resolve the problem of different image sizes in the database an input image is either increase or decrease in size. Unifying the image size will get the same number of features from all images. Moreover, resizing the image reduces processing time and thus increases system performance.

3.3 Image Feature Extraction:

In the beginning, Convolutional Neural Network (CNN) is a set of stacked layers involving both nonlinear and linear processes. These layers are learned jointly. The main building blocks of any CNN model are a convolutional layer, pooling layer, nonlinear Rectified Linear Units (ReLU) layer connected to a regular multilayer neural network called a fully connected layer, and a loss layer at the backend. CNN has known for its significant performance in applications such as visual tasks and natural language processing.

3.4 CNN Model:

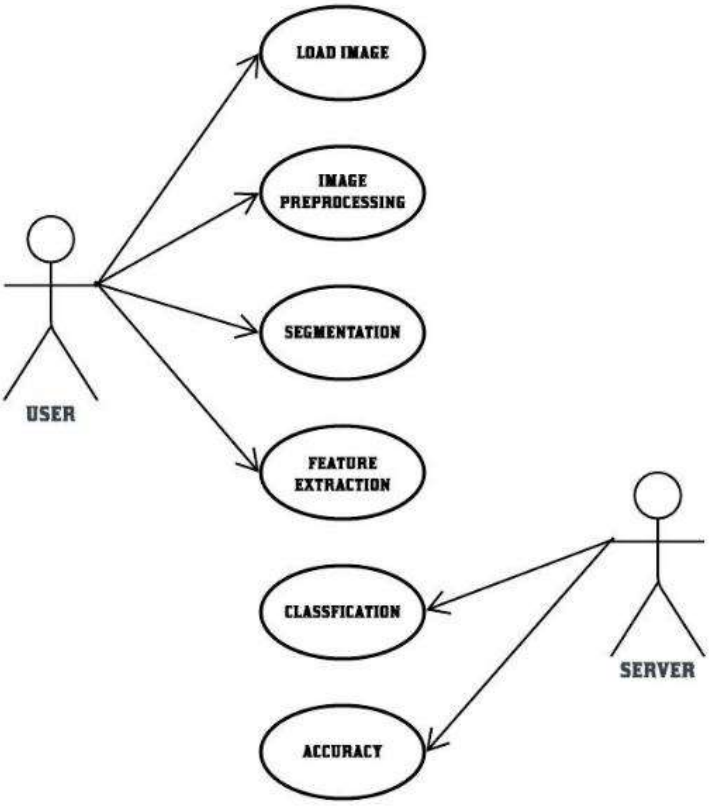
After the Extraction training model is created. training the model, the model weight and architecture can be saved with the .h5 file extension which is a Keras file. Tensorflow is needed in this process where you freeze the graph producing a .h5 file and a txt file containing the labels of the created model. To fully use the model for inference, a Tensorflow utility is used which is the Image Data Generator function and with this, the model can be successfully loaded into your application. It shows the block diagram of the necessary steps in deploying the model.

3.5 Experiments and Analysis:

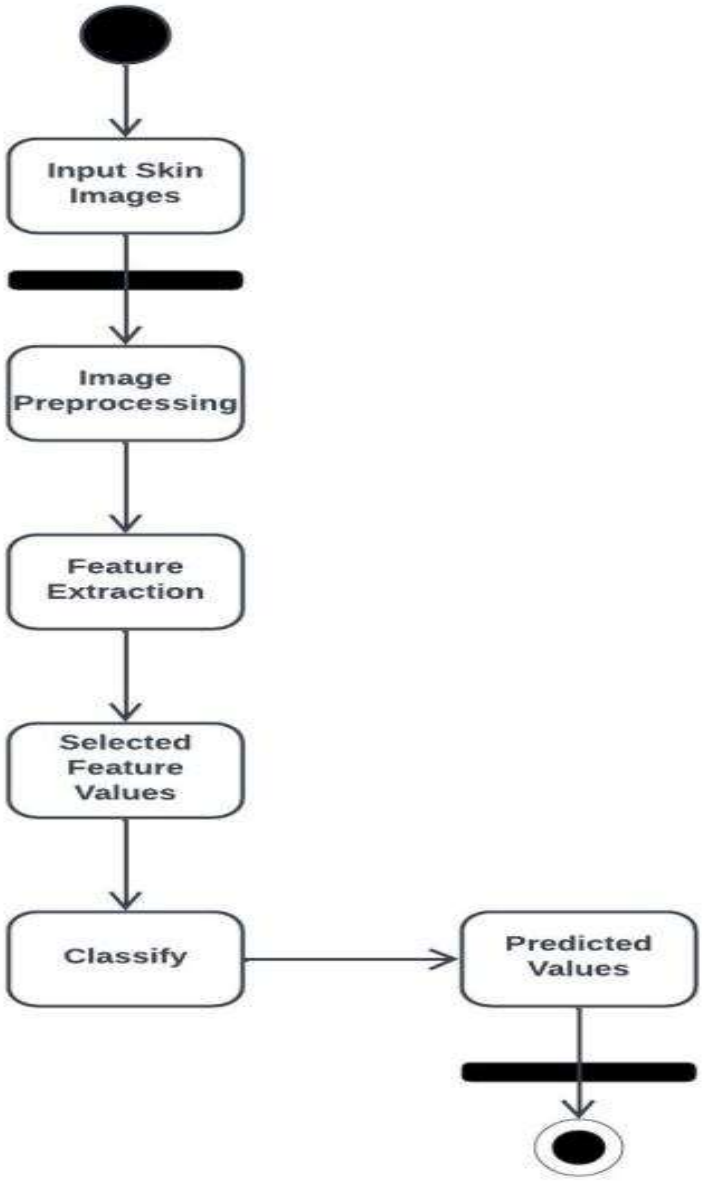
To identify the classification, in this experiment, common skin diseases (herpes, dermatitis, and psoriasis) are selected as the research objects. Ninety images are also selected to be identified accordingly, including herpes, dermatitis, and psoriasis, thirty cases in each, together with twenty test samples and ten standard samples. In this paper, the combination of the color feature and the texture feature is employed to experiment, the results of which are compared In this paper, the color feature and the texture feature can be used to make up for the weakness of the unidirectional recognition so that the recognition rate can reach to 90% and more, which greatly improves the accuracy. Ten local vertical images after segmentation can be obtained by vertical image segmentation, therefore improving the accuracy of skin disease identification.

4. UML Diagrams

4.1 Skin Cancer Use:



4.2 Activity Diagram:



5. Conclusion

Skin diseases are ranked the fourth most common cause of human illness, but many still do not consult doctors. We presented a robust and automated method for the diagnosis of dermatological diseases. Skin treatments are more effective and less disfiguring when found early. We should point out that it is to replace the human input on analysis and intuition. Research in the Indian society of medical oncology has shown for the first time that the form of Al or ML is better than experienced dermatologists of the system and the implementation methodology is presented.

References

1) Lomas, J. Leonardi-Bee, and F. Bath-Hextall, "A systematic review of the worldwide incidence of nonmelanoma skin cancer," Br J Dermatol, vol. 166, no. 5, pp. 1069–1080, 2012.A.-V. Giblin and J. M. Thomas, "Incidence, mortality and survival in cutaneous melanoma," J Plast Reconstr Aesthet Surg, vol. 60, no. 1, pp. 32–40, 2007.

2) A. V. Giblin and J. M. Thomas, "Incidence, mortality and survival in cutaneous melanoma," J Plast Reconstr Aesthet Surg, vol. 60, no. 1, pp. 32–40, 2007.

3) Z. Apalla, A. Lallas, E. Sotiriou, E. Lazaridou, and D. Ioannides, "Epidemiological trends in skin cancer," Dermatol Pract Concept, vol. 7, no. 2, pp. 1–6, 2017.

4) D. C. Whiteman, A. C. Green, and C. M. Olsen, “The growing burden of invasive melanoma: projections of incidence rates and numbers of new cases in six susceptible populations through 2031,” J Invest Dermatol, vol. 136, no. 6, pp. 1161–1171, 2016.

5) M. J. Eide et al., “Identification of patients with nonmelanoma skin cancer using health maintenance organization claims data,” Am J Epidemiol, vol. 171, no. 1, pp. 123–128, 2009.

6) C. Pouplard et al., “Risk of cancer in psoriasis: a systematic review and meta-analysis of epidemiological studies,” J Eur Acad Dermatol, vol. 27, pp. 36–46, 2013.

7) A. Egeberg, J. P. Thyssen, G. H. Gislason, and L. Skov, “Skin cancer in patients with psoriasis,” J Eur Acad Dermatol, vol. 30, no. 8, pp. 1349–1353, 2016.

8) L. Schmitz, T. Gambichler, G. Gupta, M. Stucker, and T. Dirschka, “Actinic keratosis area and severity index (AKASI) is associated with the incidence of squamous cell carcinoma,” J Eur Acad Dermatol, vol. 32, no. 5, pp. 752–756, 2018.

9) B. A. Lober and C. W. Lober, “Actinic keratosis is squamous cell carcinoma.,” South Med J, vol. 93, no. 7, pp. 650–655, 2000

10) W. C. Fix et al., “MART-1-labeled melanocyte density and distribution in actinic keratosis and squamous cell cancer in situ: Pagetoid melanocytes are a potential source of misdiagnosis as melanoma in situ,” J Cutan Pathol, vol. 45, no. 10, pp. 734–742, 2018.

11) K.-B. Tan et al., “Simulators of squamous cell carcinoma of the skin: diagnostic challenges on small biopsies and clinicopathological correlation,” J Skin Cancer, vol. 2013, 2013.

12) C. K. Bichakjian et al., "Merkel cell carcinoma: a critical review with guidelines for multidisciplinary management,” Cancer, vol. 110, no. 1, pp. 1–12, 2007.

13) M. B. Amin et al., “The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more ‘personalized’ approach to cancer staging,” CA Cancer J Clin, vol. 67, no. 2, pp. 93–99, 2017.

14) R. S. Padilla, S. Sebastian, Z. Jiang, I. Nindl, and R. Larson, “Gene expression patterns of normal human skin, actinic keratosis, and squamous cell carcinoma: a spectrum of disease progression,” Arch Dermatol, vol. 146, no. 3, pp. 288–293, 2010.

15) Z. Wang, M. Gerstein, and M. Snyder, “RNA-Seq: A revolutionary tool for transcriptomics,” Nat Rev Genet, vol. 10, no. 1, pp. 57–63, 2009.