As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 9(05): 833-838, 2023
Copyright © 2023International Journal for Modern Trends in Science and Technology
ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0905142
Available online at: http://www.ijmtst.com/vol9issue05.html

Skin Disease Detection using Machine Learning

G. Sindhura Bhargavi, D. Venkata Sai, K. Likith Reddy, A. Naveen, E. Rajesh

Department of Electronics and Communication Engineering, Narayana Engineering College, Nellore, Andhra Pradesh, India

To Cite this Article
G. Sindhura Bhargavi, D. Venkata Sai, K. Likith Reddy, A. Naveen, E. Rajesh. Skin Disease Detection using Machine Learning. International Journal for Modern Trends in Science and Technology 2023, 9(05), pp. 833-838. https://doi.org/10.46501/IJMTST0905142

Article Info
Received: 21 April 2023; Accepted: 20 May 2023; Published: 24 May 2023.

ABSTRACT

Skin diseases are more common than other diseases. Skin diseases may be caused by fungal infection, bacteria, allergy, or viruses, etc. The advancement of lasers and Photonics based medical technology has made it possible to diagnose the skin diseases much more quickly and accurately. But the cost of such diagnosis is still limited and very expensive. So, image processing techniques help to build automated screening system for dermatology at an initial stage. The extraction of features plays a key role in helping to classify skin diseases. Computer vision has a role in the detection of skin diseases in a variety of techniques. Due to deserts and hot weather, skin diseases are common in Saudi Arabia. This work contributes in the research of skin disease detection. We proposed an image processing-based method to detect skin diseases. This method takes the digital image of disease effect skin area, then use image analysis to identify the type of disease. Our proposed approach is simple, fast and does not require expensive equipment other than a camera and a computer. The approach works on the inputs of a colour image. Then resize the of the image to extract features using pretrained convolutional neural network. After that classified feature using Multiclass SVM. Finally, the results are shown to the user, including the type of disease, spread, and severity.

KEYWORDS: Dermatoscopic images, Deep Learning, Data Enhancement, Convolutional Neural Network(CNN), Model Training, Testing and Evaluation.

1. INTRODUCTION

Artificial Intelligence is taking over automation in all fields of application even within the healthcare field. In the past years these diseases have been a matter of concern due to the sudden arrival and the complexities which has increased life risks. These Skin abnormalities are very infectious and the require to be treated at earlier stages to avoid it from spreading. The majority of diseases is caused by unprotected exposure to excessive Ultraviolet Radiation(UR). Among all, benign type is considered to be less dangerous than malignant melanoma and can be cured with proper treatment, whereas the deadliest form of skin lesion is malignant Melanoma. The survey results indicate that the back and lower extremity, trunk and upper extremity are heavily compromised regions of skin cancer. There are large instances of patients with age ranging from 30 to 60. Also, MelanocyticNevi, Carcinoma and Dermatofibroma are not prevalent below the age of 20years.

2. EXISTINGTECHOLOGY

A. Artifical Neural Network(ANN)

An artificial neuron network (ANN) is a statistical nonlinear predictive modelling method which is used to learn the complex relationships between input and

output. The structure of ANN is inspired by the biological pattern of our brain neuron [2]. An ANN has three types of computation node. ANNs learn computation at each node through back-propagation. There are two sorts of data set trained and untrained data set which produces the accuracy by employing a supervised and unsupervised learning approach with different sort of neural network architectures like feed forward, back propagation method which uses the info set at a special manner. Using Artificial Neural Network, accuracy obtained in various researches is 80% which isn't optimum [2]. Also, ANNs require processors with parallel processing power. ANN produces a probing solution it does not give a clue as to why and how it takes place which reduces trust in the network.

B. Back Propagation Network(BPN)

Back propagation, a strategy in Artificial Neural Networks to figure out the error contribution of each neuron after a cluster of information (in image recognition, multiple images) is processed. Back Propagation is quite sensitive to noisy and uproarious data. The BNN classifier achieves 75%-80% accuracy [2]. BNN is benefits on prediction and classification but the processing speed is slower compared to other learning algorithms [5] [2]. *C . Support Vector Machine(SVM).*

SVM is a supervised non-linear classifier which constructs an optimal n-dimensional hyperplane to separate all the data points in two categories [2]. In SVM, choosing an honest kernel function isn't easy. It requires long training time for large datasets. Since the final model is not easy to use we cannot make small calibrations to the model and it becomes difficult to tune the parameters used in SVMs.

SVMs when compared with ANNs always give best results [3].

3. LITERATURE

Skin diseases are the 4th common cause of skin burden worldwide. Robust and Automated system have been developed to lessen this burden and to help the patients to conduct the early assessment of the skin lesion. Mostly this system available in the literature only provide skin cancer classification. Treatments for skin are more effective and less disfiguring when found early and it is a challenging research due to similar characteristics of skin diseases. In this project we attempt to detect skin diseases .A novel system is presented in this research work for the diagnosis of the most common skin lesions (Melanocytic nevi, Melanoma, Benign keratosis-like lesions, Basal cell carcinoma, Actinic keratoses, Vascular lesion, Dermatofibroma). The proposed approach is based on the pre-processing, Deep learning algorithm, training the model , validation and classification phase. Experiments were performed on 10010 images and 93% accuracy is achieved for seven-class classification using Convolution Neural Networks (CNN) with the Keras Application API.

4. DATASET

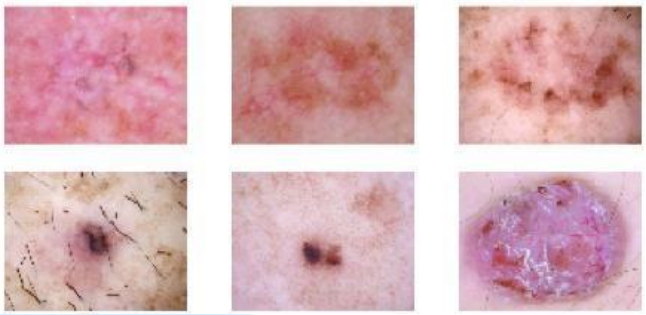

Fig. 1. Sample Data

The fig. above is the sample dataset which we have trained and tested

5. IMPLEMENTATION (METHODOLOGY)

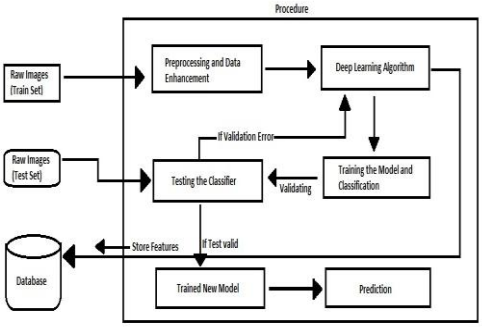

Fig.2. Procedure

To develop any ML-AI based system, be it this system; following steps are to be followed.

1. Data Gathering

The proposed system has been assessed on dermatoscopic images which is collected from

publicly available dataset based on SkinCancer-MNIST (Modified National Institute of Standards and Technology Database)-HAM10000. The number of options is endless. To save time and effort one can use publicly available data.

2. Data Preprocessing &Enhancement

"Trash In- Good Out" is the basic motto in this step [6]. Validating your dataset with some basic profiling procedure will help speeding up the process, by slip-ups and grimy information [4]. AI algorithms don't give great outcomes when working with such information.

2.1. Data Cleaning.

Dirty data can cause confusion and results in unreliable and poor output. Hence first step in Data Pre-processing is Data Cleaning. Cleaning of data is done by filling in missing values, smoothing noisy data by identifying and/or removing outliers, and removing inconsistencies.

2.2. Data Transformation.

Data Transformation involves converting data from one format into another. It involves transforming actual values from one representation to the target representation.

2.3. Exploratory Data Analysis (EDA). In this we explore different features of the dataset, their distributions and actual counts.

2.4. Label Encoding.

The dataset is labelled into 7 different categories:

1. MelanocyticNevi
2. Melanoma
3. Benign keratosis-like lesions
4. Basal cell carcinoma
5. ActinicKeratoses
6. Vascular lesions 7. Dermatofibroma

3. Training.

For this we have to divide the data into training set and testing set. This division can be in any ratio. Also, the batch size and number of epochs has to be decided beforehand.

4. Model Building

We have used Convolutional Neural Network (CNN). A **Convolutional Neural Network** (**CNN** or **ConvNet**) is a category of deep neural networks, where the machine learns on its own and divide the data provided into the levels of prediction and in a very short period of time gives the accurate results [2]. A Convolutional Neural Network (CNN) is an algorithm in deep learning which consist of a combination of convolutional and pooling layers in sequence and then followed by fully connected layers at the end as like multilayer neural network [2]. CNN stands out among all alternative algorithms in classifying images. Crucial characteristics are Sparse Connectivity, Shared Weights and Pooling Feature so as to extract the best features. Also, the use of Graphical Processing Units (GPUs) have shrivelled the training time of deep learning methods. Giant databases of lasbelled data and pre-trained networks are now publicly available.

The figure below shows the difference between Sparse and Dense Connectivity.

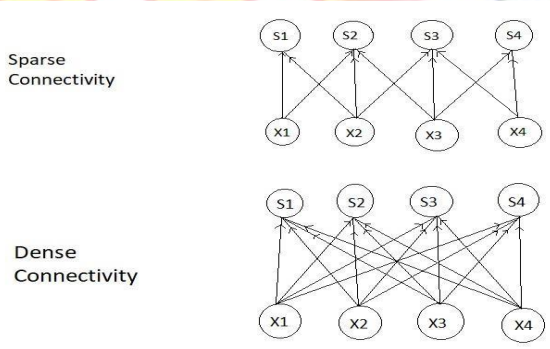

Fig.3. Sparse and Dense Connectivity

4.1. Explanation

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. Next important layer is the pooling layer which simply acts as a down sampling filter. We have Max pooling, MaxPool() picks the **maximal worth** among set **of** two neighbouring pixels. This layer is employed to scale back (to cut back) machine value and additionally reduce overfitting to some extent. 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 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 in generalizing the network. Now, to convert the final feature maps into a 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 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 actually measure the error rate between the observed labels and predicted labels. Next most important is the optimizer. **Adam Optimizer** has 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 rate so as to have minimal loss. **ReduceLROnPlateau** is used, the name itself means reduce the LR so as to reach the global minimum of loss function.

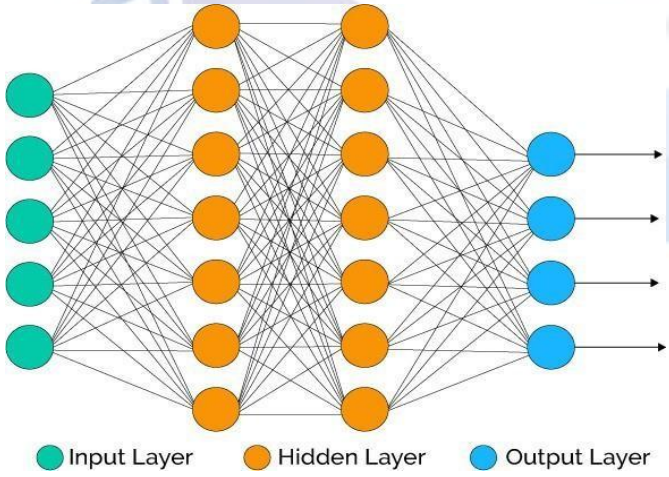

Fig.4. Fully Connected Network

5. Model Evaluation.

"More the accuracy, better is the model". Every model is evaluated based on the accuracy achieved and the loss obtained. There are two accuracies involved: Validation accuracy And Test accuracy. Before this Validation set is different from Train set i.e. Validation set is independent from the Train set, Validation set is used for selecting parameters. Just for an instance if your model has 90% train accuracy and 89% validation accuracy then your model is expected to have 89%accuracy on new data.

6. Graphical Analysis.

This involves plotting Histogram and the Confusion Matrix.

Confusion Matrix involves TP, TN. FP, FN [2]. The set of decision made by classification algorithm contains correct decision (true positive) and incorrect decision (false positive). False negative are those decision which is declared as negative by classification algorithm but actually they are positive. True negative are those decision which is correctly identified as negative by the classification algorithm.

6. OVERVIEW

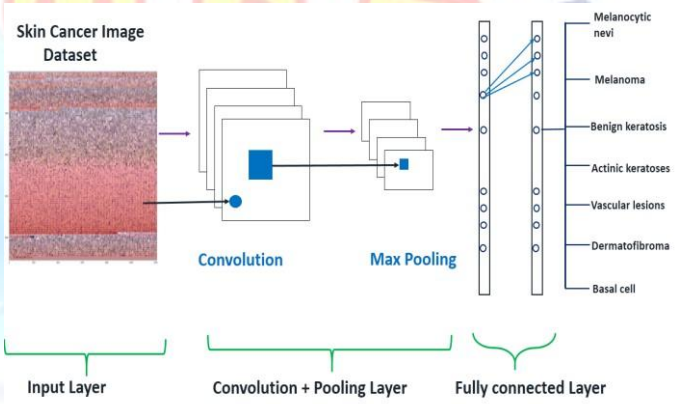

Fig.5. Convolution Neural Network (CNN)

7. RESULTS

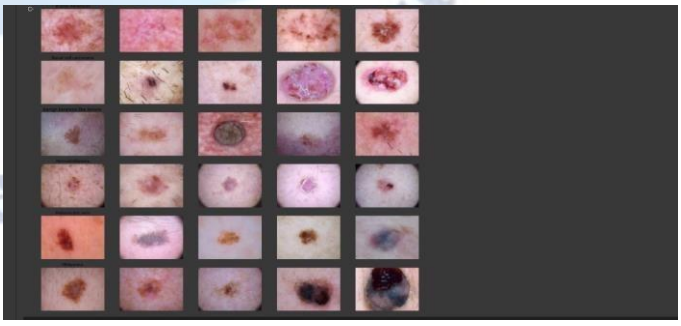

Fig.6. Streamed Images

Fig.7. Model Summary

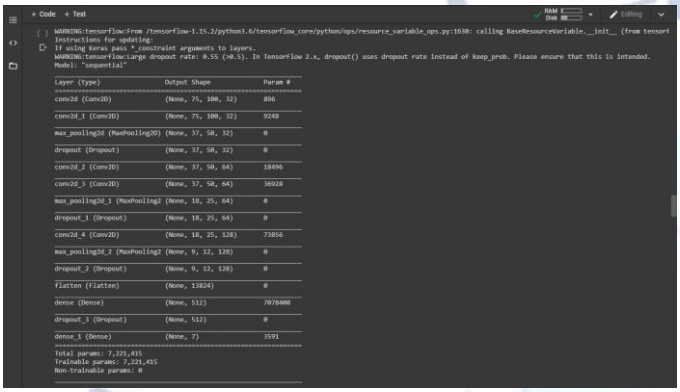

Fig.8. Epoch-50

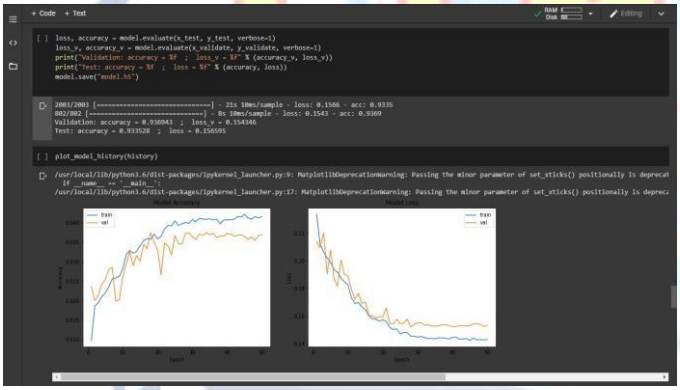

Fig.9. Graphical Plotting for Epoch-50

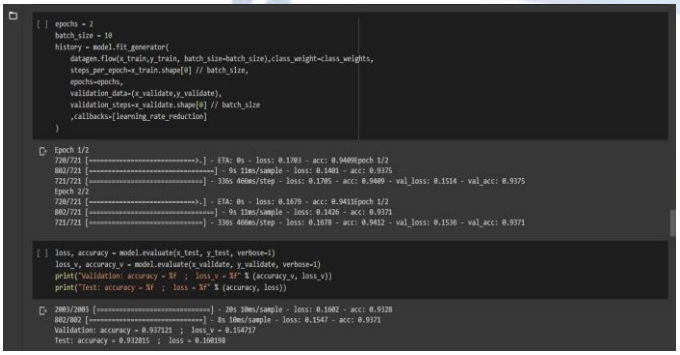

Fig.10. Epoch-2

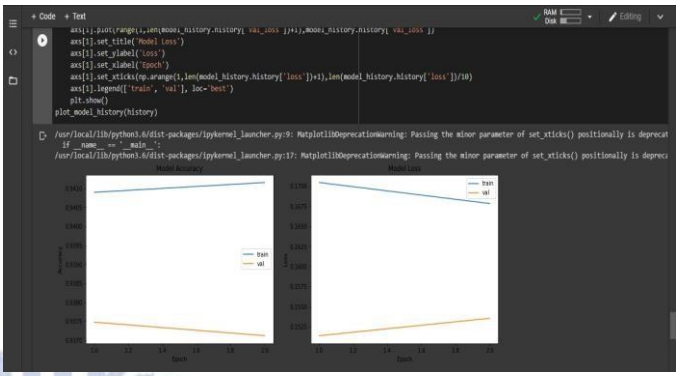

Fig.11. Graphical Plotting for Epoch-2

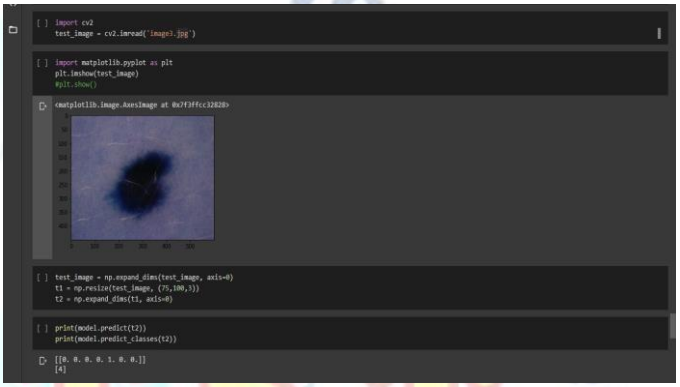

Fig.12. Testing for an Image with Detection

EPOCH-1 TABLE I.

Sr. No.	Evaluation		
	Metric/Parameter	***Testing***	***Validation***
1.	Accuracy	93.35%	93.35%
2.	Loss	15.65%	15.65%

TABLE II.

EPOCH-2

Sr. No.	Evaluation		
	Metric/Parameter	***Testing***	***Validation***
1.	Accuracy	93.28%	93.28%
2.	Loss	16.01%	16.01%

8. CONCLUSION

Skin Diseases are ranked 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. Treatments for skin are more effective and less disfiguring when Skin Diseases are ranked fourth most common cause of human illness, found early. We should point out that it is to replace doctors because no machine can yet replace the human input on analysis and intuition. Researches in European Society of Medical Oncology have shown for the first time that form of AI or ML is better than experienced dermatologists. In this a brief description of the system and the implementation methodology is presented.

FUTURE SCOPE

Once improvements are made on the Inception V3 algorithms for the skin cancer in the field of medical the results can be improved. So nowadays deep learning is the best solution of skin cancer diseases classification and recognition of cancerous diseases. In the future Inception V3 for skin cancer disease would help medical stakeholders to avoid the conventional lab and in-vivo tests. Without testing or without the use of x rays skin cancer would be automatically detected using the Inception V3 algorithm embedded in mobile. This saves time; its simplicity makes it easy for the public without the need of a doctor.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., ... &Utikal, J. S. (2019). Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. European Journal of Cancer, 113, 47-54.

[2] Bi, L., Kim, J., Ahn, E., & Feng, D. (2017). Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks. arXiv preprint arXiv:1703.04197.

[3] Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., & Smith, J. R. (2015, October). Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopyimages.In International workshop on machine learning in medical imaging (pp. 118-126).

[4] Damsky, W. E., &Bosenberg, M. (2017). Melanocytic nevi and melanoma: unraveling a complex relationship. Oncogene, 36(42), 5771.

[5] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., &Thrun,S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115.

[6] Grayson, W., &Pantanowitz, L. (2008). Histological variants of cutaneous Kaposi sarcoma. Diagnostic pathology, 3(1), 31.