iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 Issue: VI Month of publication: June 2022

DOI: https://doi.org/10.22214/ijraset.2022.43872

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Skin Cancer Detection with the Aid of Deep Learning

Badarinath H R[1], Darshan G[2], Pramod K L[3], Prasanna R[4], Prof. Manjesh[5]
[1, 2, 3, 4]Department of Information Science and Engineering, Jyothy Institute of Technology
[5]Guide

***Abstract:** Skin cancer is now regarded as one of the most dangerous types of cancer seen in humans. Clinical screening is followed by dermoscopic analysis and histological testing in the diagnosis of melanoma. Melanoma is a type of skin cancer that is highly treatable if caught early. Effective segmentation of skin lesions in dermoscopy pictures can increase skin disease categorization accuracy, giving dermatologists a powerful tool for studying pigmented skin lesions. The goal of the research is to create an automated classification system for skin cancer utilising photos of skin lesions that is based on image processing techniques. Deep Learning models embed different neural networks, such as Convolutional Neural Networks (CNN), which are well-known for capturing spatial and temporal correlations. with the use of appropriate filters in an image Individual transformational aspects that are limited by the data augmentation procedure derive useful and particular data for training the algorithm to make attractive predictions.*

***Index Terms:** Convolution Neural Network(CNN), Melanoma, Feature Extraction, ABCD of Skin Cancer*

I. INTRODUCTION

Melanoma is the least frequent type of skin cancer, but it is the most deadly, accounting for 75 percent of all deaths. Although this is a less common type of skin cancer, it can swiftly spread to other places of the body if not detected early. Perusing the statistical data provided by the World Health Organization, 13.1 million sufferers of melanoma can be seen by the end of 2030. The primary cause of melanoma, as per the World Cancer Report, is ultraviolet light exposure in those with a low degree of skin pigment. UV rays can come via solar rays or just about any source, and moles account for about 25% of malignant moles.

The furious advancement in the fields of image processing and classification availed to adopt the new tools and technologies to bring the CNN process into medical diagnosing domains. The underlying neural network expands from the input to the output layer with the vision of mapping the functions to the expected layer with a bit of tuning to network parameters and a goal to attain maximum accuracy. CNN possesses standardized layers. The initial layer accounts for the data collected by dermatologists and accepts it as input. Further, the information is grouped and transported to the pooling layer. The pooling layer performs the structurization of the data based on max and min pool functionalities. The smoothing process includes the straightening of the layer which converts the available information into a One-Dimensional vector. Class conversion happens in a thick layer where the information is biased into benign and malignant. The process depicts the automated way of analyzing skin cancer and classifies it as benign or malignant melanoma. We solve the challenge by creating deep learning-based image classification models for skin cancer detection without any prior programming experience.

A. Objectives

1) Redistribution of collected datasets based on the seven types of skin cancers.
2) Feature extraction involved in Image processing follows the ABCDE motto of skin cancer synthesis.
3) Data pre-processing involves the statistical methodologies to train the model in an efficient manner.
4) Deploying the rethinking models to maximize the scaling and efficient use of resources as well as to increase the predictability of the software.

B. Deep Learning

Deep Learning methodology imitates the human brain accordingly with the aid of entities like data inputs, weights, and bias, which drive the accuracy and assist in recognition and classification of objects residing in the data. To refine the optimization and prediction, the deep learning network works with the nodes that are interconnected to form multiple layers. Acyclic formation of nodes that depicts the **Forward propagation** analogy. The visible layer constitutes the initial input and final output layers that enable the user to observe the variables. The input layer engulfs the required data for processing, whereas the output layer performs the prediction and classification of the processed data.

Backpropagation is an algorithm-driven process that trains the model using a gradient of loss function and errors are calculated whilst the prediction. Traversing backwards through the layers it will reduce the redundancy.

To increase the accuracy of the algorithm and to avoid the errors during prediction process of a neural network, both forward and backward propagation methodologies are adopted.

C. Application of CNN in Image processing Domain

The neurons housed in our brains are the motivation to build these Artificial Neural Networks. The generally used feed-forward neural network for image classification is slightly modified with the involvement of CNN. Translational invariance can only be defined under CNN where it detects the objects in their distinct shapes, which gives a trump card for CNN over the feed-forward neural network. According to Layman, the feed-forward neural network's conventional approach attempts to find the details of the object if it lies at the center of the image. It would fail if the object lies off the center or if any slight positional changes would distract the network to detect the object. This leads to an inconvenient way of processing the data.

D. Classification Of Skin Lesions Using CNN

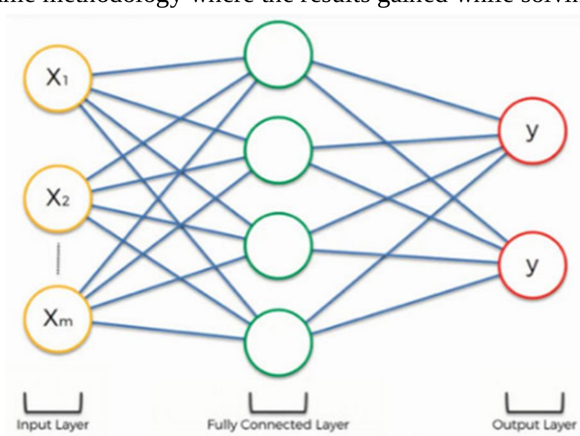
In the epoch, the comparative conclusion between the human and machine, the triumph of the machine is praisable. The dermatologists account for more time-consuming than the deep learning model.

The two distinct ways followed by CNN to conclude the classification process.

1) *Feature Extraction of Images:* To attain better results it converts the raw data into numerical one without disturbing the originality.
2) *Model Deployed will perform the classification:* End-to-end processing of the skin lesion images to extract the results and classify them into 7 classes of skin cancer.

This process requires a model that is pre-trained (where the large set of images are fed to the model for the qualified output or to increase the learning capability) or it must be built from scratch which doesn't seem to be feasible.

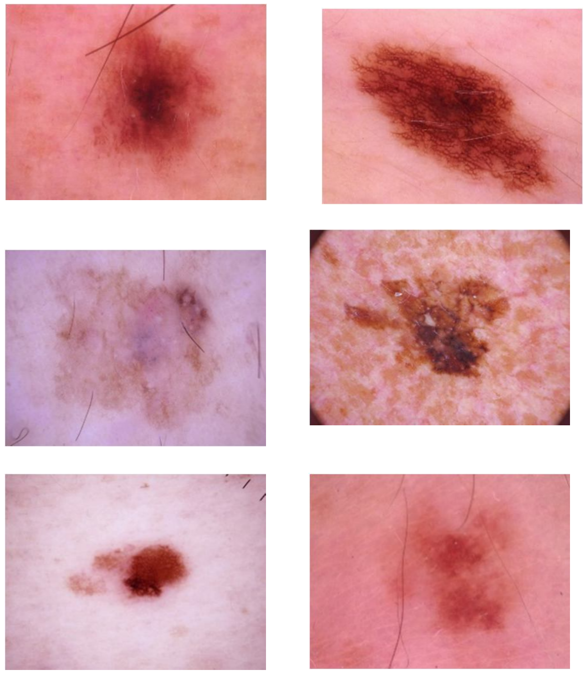
Transferred Learning deals with the same methodology where the results gained while solving a problem are applied to the other.



E. Datasets

1) Human Against Machine (HAM 10000)

Comparison between the accuracy of human observed results and the convoluted artificial networks. The revolution in AI makes CNN decide almost like the humans. In our approach, the clinical images were collected without a dermatoscopic diagnosis to check the accuracy of the model. The challenge of collecting diverse images to train neural networks is tackled by HAM which includes the images of skin lesions from the diverse population and stored by different modalities. Further semi-automated workflows are developed to deal with the datasets like cleaning and acquisition. The credits go to ISIC, which gives the dataset for the academic research in machine learning which constitutes 10015 images. Provided the benchmark to achieve the decisional capability of humans, the diagnosis has to qualify the accuracy compared to dermatoscopic results with human intervention.

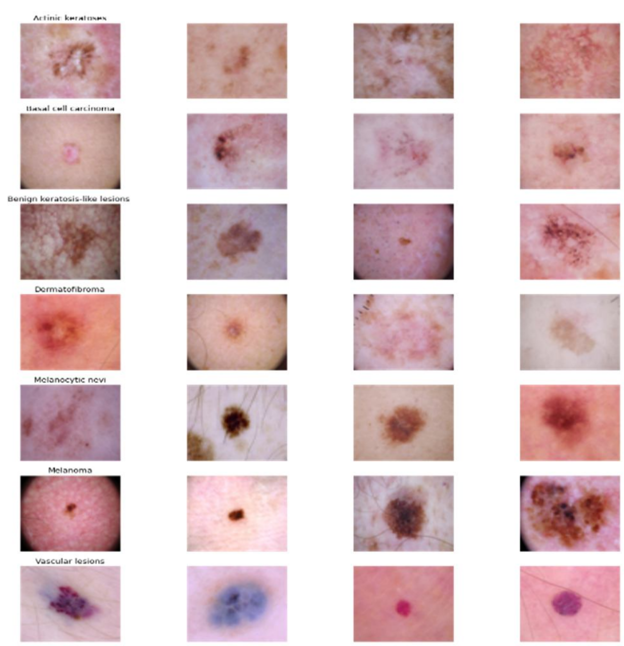


II. WORKING OF MODEL

A. Classification of the Dataset

Displaying the datasets according to the classified index as distinct “7types of skin cancer”.

Melanocytic nevi (nv)
Benign keratosis-like lesions (bkl)
Melanoma (mel)
Basal cell carcinoma (bcc)
Actinic keratoses (akiec)
Vascular lesions (vas)
Dermatofibroma (df)



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

B. Resizing the Dataset

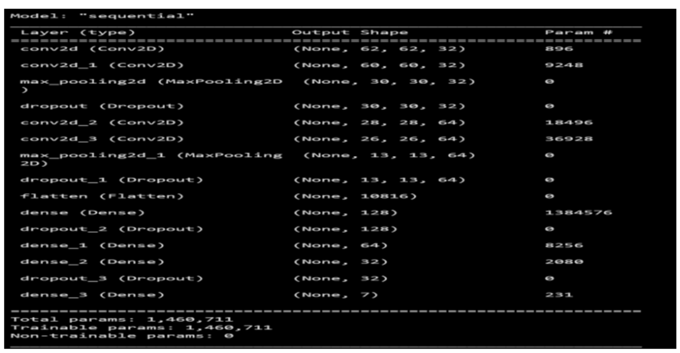
1) Data Augmentation

There is a need to generate new data points from the collected one and the techniques to do the same with the aid of deep learning model are defined under data augmentation process.
The property called invariance which is defined under the convolutional neural network helps to classify the objects in different orientations quickly as possible.
The importance of data augmentation can be seen while working with a dataset that has a limited set of conditions. The process evolves the dataset with a myriad variety such as measurements of pixels, positioning, tuning of the images, etc.

	lesion_id	image_id	dx	dx_type	age	sex	localization	path	cell_type	cell_codes	image
0	HAM_0000118	ISIC_0027419	bkl	histo	80.0	male	scalp	archive/ham10000_images_part_1/ISIC_0027419.jpg	Benign keratosis-like lesions	2	[[[191, 152, 194], [191, 153, 195], [192, 149,...
1	HAM_0000118	ISIC_0025030	bkl	histo	80.0	male	scalp	archive/ham10000_images_part_1/ISIC_0025030.jpg	Benign keratosis-like lesions	2	[[[23, 13, 23], [25, 14, 28], [37, 24, 46], [6...
2	HAM_0002730	ISIC_0026769	bkl	histo	80.0	male	scalp	archive/ham10000_images_part_1/ISIC_0026769.jpg	Benign keratosis-like lesions	2	[[[186, 129, 140], [192, 136, 151], [198, 143,...
3	HAM_0002730	ISIC_0025661	bkl	histo	80.0	male	scalp	archive/ham10000_images_part_1/ISIC_0025661.jpg	Benign keratosis-like lesions	2	[[[24, 11, 19], [36, 20, 30], [64, 38, 50], [9...
4	HAM_0001466	ISIC_0031633	bkl	histo	75.0	male	ear	archive/HAM10000_images_part_2/ISIC_0031633.jpg	Benign keratosis-like lesions	2	[[[139, 95, 118], [158, 114, 138], [178, 133, ...

2) CNN Model

The Convolutional Neural Networks deal with the processing techniques in a distinct way. It includes 3 layers. i.e, Input, Output and Hidden layers
In this model, we first implement new feature extractor or kernels to the source image to produce the first two convolved layers, then we apply Max Pooling to gain the max pooled layer, then we obtain the next layers in the same manner, and ultimately we acquire the Flatten layer using Flatten (). The completely linked layers (hidden layers) are then added, culminating in a 7-neuron output layer.

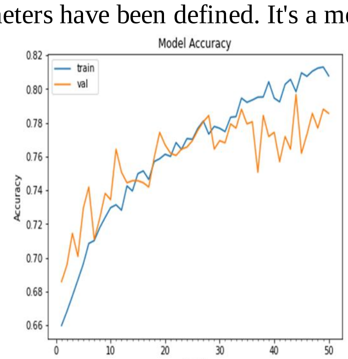


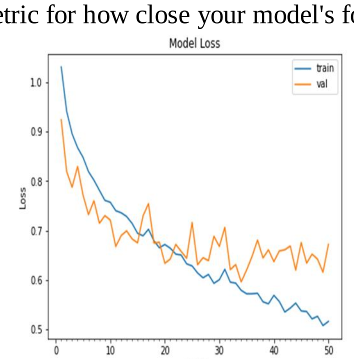
```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 62, 62, 32)        896
 conv2d_1 (Conv2D)           (None, 60, 60, 32)        9248
 max_pooling2d (MaxPooling2D  (None, 30, 30, 32)       0
 )
 dropout (Dropout)           (None, 30, 30, 32)        0
 conv2d_2 (Conv2D)           (None, 28, 28, 64)        18496
 conv2d_3 (Conv2D)           (None, 26, 26, 64)        36928
 max_pooling2d_1 (MaxPooling  (None, 13, 13, 64)       0
 2D)
 dropout_1 (Dropout)         (None, 13, 13, 64)        0
 flatten (Flatten)           (None, 10816)             0
 dense (Dense)               (None, 128)               1384576
 dropout_2 (Dropout)         (None, 128)               0
 dense_1 (Dense)             (None, 64)                8256
 dense_2 (Dense)             (None, 32)                2080
 dropout_3 (Dropout)         (None, 32)                0
 dense_3 (Dense)             (None, 7)                 231
=================================================================
Total params: 1,460,711
Trainable params: 1,460,711
Non-trainable params: 0
_________________________________________________________________
```

C. Performance Evaluation

1) Validation loss and Validation Accuracy Graph

A machine learning algorithm is optimised using a loss function. The loss is calculated using training and validation data, and its meaning is determined by how well the model performs in these two sets. It's the total number of errors committed in each training or validation set for each example. The loss value indicates how well or poorly a model performs after each optimization iteration.
An accuracy metric is used to interpretably measure the algorithm's performance. The accuracy of a model is usually calculated as a percentage after the model parameters have been defined. It's a metric for how close your model's forecast is to the actual data.

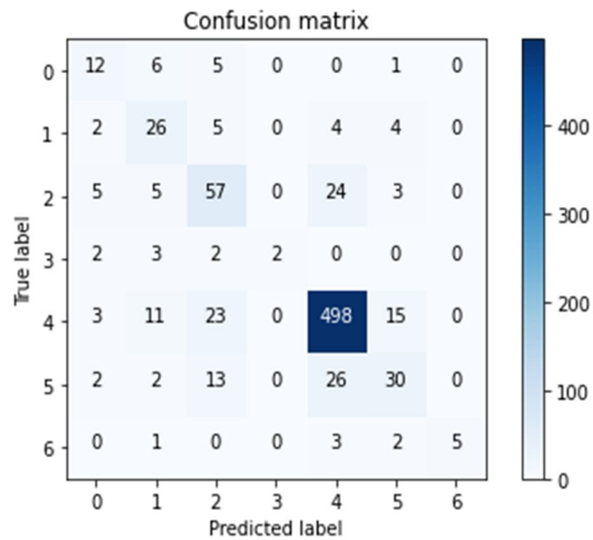




International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

2) Confusion Matrix

A confusion matrix is a table that shows how many correct and incorrect predictions a classifier made. It's a metric for evaluating a classification model's performance. It can be used to calculate performance metrics like accuracy, precision, recall, and F1-score to evaluate the performance of a classification model.



III. CONCLUSION

The depicted model concludes whether the given dataset is melanoma or not. Further with 78% of accuracy it will provide assistance in diagnosis domain. It aims in developing a user friendly GUI which helps the clients to obtain the results in fraction of second. The robustness of the model is enhanced over the time.

REFERENCES

[1] Xie F., Fan H., Li Y., Jiang Z., Meng R., Bovik A. Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model. IEEE Trans. Med. Imaging. 2017;36:849–858. DOI: 10.1109/TMI.2016.2633551. [PubMed] [CrossRef] [Google Scholar]

[2] G.P. Guy, C.C. Thomas, T. Thompson, M. Watson, G.M. Massetti, L.C. Richardson, Vital signs: melanoma incidence and mortality trends and projections—United States, 1982–2030. Morb. Mortal. Wkly. Rep. (2015)

[3] “World Cancer Research Fund - Skin Cancer Statistics,” 2018. [Online]. Available: https:// www.wcrf.org/dietandcancer/cancer-trends/skin-cancer-statistics. Accessed: 28-Oct-2019

[4] American Cancer Society, “Cancer Facts and Figures 2019.” [Online]. Available: https://www. cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-factsand-figures/2019/cancer-facts-and-figures-2019.pdf. Accessed: 29-Oct-2019

[5] K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis, Machine learning applications in cancer prognosis and predictio. Comput. Struct. Biotechnol. J. (2015)

[6] Step wise Approach: CNN Model. (2019, July, 20) [Online]. Available: https://www.kaggle.com/sid321axn/step-wise-approach-cnn-model-77-0 344-accuracy

[7] ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection. (2019, July, 25) [Online]. Available: https://challenge2018.isic-archive.com/

[8] Skin Cancer MNIST: HAM10000. (2019, July, 27)[Online].Available:https://www.kaggle.com/kmader/skin-cancer-mnist-ham10000

[9] Z. Gao et al., HEp-2 cell image classification with deep convolutional neural networks. IEEE J. Biomed. Health Inf. 21(2), 416–428 (2016)

[10] B. Harangi, Skin lesion classification with ensembles of deep convolutional neural networks. J. Biomed. Inform. 86, 25–32 (2018)