International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252

Prediction of Skin Cancer using Machine Learning

Anurag Hans
(Student, B.Tech IT, MAIT, New Delhi)

Submitted: 05-06-2021 Revised: 18-06-2021 Accepted: 20-06-2021

ABSTRACT
Cancer is a deadly disease that occurs by multiplication of the cells in an uncontrolled manner and thereby occupying the peripheral tissues. Although, skin cancer is less frequent than many other cancer types, itsmortality rate is quite high. It is of different types such as Malignant Melanoma, Basal cell carcinoma and Squamous cell carcinoma. The mortality rates from melanoma skin cancer have risen by 156% in the UK. Skin cancer is primarily diagnosed visually, beginning with an initial clinical screening by a dermatologist, and potentially followed by dermoscopic analysis, a biopsy, and histopathological examination and these processes easily take a month. According to WHO, early detection increases the chances for successful treatment. So therefore the model using Machine Learning algorithms detecting cancer quickly will help in diagnosing skin cancer.

I. INTRODUCTION

Cancer is a disease that occurs through multiplying of the body cells in an uncontrolled manner and occupying the peripheral tissues. Although the skin cancer disease occurs less frequent than many other cancer types, it is highly important because of its high mortality. The skin cancer is of different types, such as Malignant Melanoma, Squamous cell carcinoma, and Basal cell carcinoma. Melanoma incidence is being reported to increase more rapidly than the other forms of cancer. Melanoma is responsible for 4% only of all skin cancers, whereas it is responsible for 75% total of skin cancer deaths.

Various research works have been done using Image processing and Computer vision to classify skin cancer into seven types. There were several limitation of previous models like some used PH2 data sets [4][5], which has only 200 dermoscopy images and some used images of lower resolution. Most of the model use KNN model [1][4][6][7][9].

Compared to the previous researches we used ISIC_MSK-2 data sets having 1535 images and we used higher resolution images with CNN model for classification of cancer.

Objectives
To study about skin cancer and related work.
To propose a model for prediction of skin cancer using high resolution images.
To automate the detection approach using faster and more accurate image classification algorithms.
To decrease treatment time with increase in prediction accuracy.

II. LITERATURE REVIEW

In Paper [1] authors proposed segmentation & classification of skin cancer Melanoma from skin lesion images. Pre-processing was done like hair removal using Dull Razor tool, after that image segmentation is done by semi-supervised mean shift algorithm, after that feature extraction is done using ABCD rule (A – Asymmetry B – Border C – Color variation D – Diameter), next step is classification, they used three classification algorithms - KNN, Decision Tree, SVM. The highest accuracy was 78.2% using SVM algorithm.

In Paper [2] authors proposed model on Image processing & SVM classification for Melanoma Detection. In this, ABCD rule was used for detection and classification. In pre – processing, changing the contrast which makes the patch looks bright, then using HSV (Hue, Saturation and value). After that segmentation phase comes in which grabcut technique was used. Accuracy of this model was 80.00%. For improving the accuracy SVM(Support Vector Machine) was used.

In Paper [3] authors proposed Melanoma detection in Dermoscopy Images using Image Processing and machine Learning work was divided into Lesion segmentation, Feature segmentation, Feature Generation and classification. In Lesion segmentation infected skin was taken for processing. Next was feature segmentation after this step the last step was feature generation and classification in which lesion was classified into the appropriate category of cancer.

In Paper [4] authors presented Skin lesion classification using machine learning algorithms uses PH2 data sets, which has 200 dermoscopy images at 768x560 resolution with each image having RGB channel. Four classifying techniques were applied on the data sets which were Artificial Neural Network(ANN), SVM, K-Nearest Neighbor(KNN), Decision Tree(DT).

In Paper [5] authors proposed Efficient Machine Learning Approach for the Detection of Melanoma using Dermoscopic Images mainly used two techniques on PH2 datasets which were feature extraction and classification, In feature extraction colour feature and texture feature were used. 13-D feature vector was formed including nine color features and four texture features. 13-D feature vector was generated for each dermoscopic image and saved into the database with their respective class labels. In the proposed method, SVM classifier was chosen to apply in order to classify melanoma images out of all given dermoscopic images based upon extracted color and texture features.

In Paper [6] authors proposed a model for the segmentation purpose, using basic algorithm together with a fusion strategy. Then KNN classification was used to classify the skin lesion into three categories that were benign, dysplastic and malignant melanoma. In order to improve the accuracy, some other measures like age, amount of lesions, sunburns etc. that indicate the general risk were also required . Proposed methodology mainly consists of four major steps namely Preprocessing(used to enhance the image and to remove the various artifacts), segmentation(Segmentation is used to find the region of interest), feature extraction(Color features play a very important role in the melanoma skin cancer diagnosis. The accuracy of the system was affected by the features used), and classification(3 Classifiers were used namely Naive Bayes , Decision Tree , K-Nearest Neighbor). MED-NODE dataset was used to evaluate the system performance for the melanoma skin cancer detection system. Accuracy is 82.35%.

In Paper [7] authors proposed a computer-aided classification of skin cancer images was an active area of research and proposed different classification methods. To deal with the problem of limited labeled data availability they presented a semi advised learning model for automated recognition of skin cancer using histo-pathalogical images. Deep belief architecture was constructed using unlabeled data by making efficient use of limited labeled data for fine-tuning done the classification model. In parallel an advised SVM algorithm was used to enhance classification results by counteracting the effect of misclassified data using advised weights. To increase generalization capability of the model, advised SVM and Deep belief network were trained in parallel. Then the results are aggregated using least square estimation weighting. The classification performance is compared with some popular methods and the proposed model outperformed most of the popular techniques including KNN, ANN, SVM and semi supervised algorithms like Expectation maximization algorithm and transductive SVM based classification model.

In Paper [8] authors proposed a novel scheme for early detection of melanoma using Multiclass support vector machine (MSVM) .They used an automatic procedure, where the queried images were grouped and matched with higher probability type to classify the type of melanoma .In the MSVM classifier algorithm, the test samples were mapped with training samples and the probability value was set for the highest match obtained group of training samples. They made a proposed system that contains an image database which has the all five types of melanoma for testing and classification purposes .The simulation results have shown the superior performance and accuracy of One-Against-All MCSVM. The accuracy of the proposed support vector machine scheme has comparatively high among all five types. To avoid the lack in accuracy, segmentation algorithm called K-means clustering algorithm is used .

In Paper [9] authors proposed a semi supervised, self-advised learning model for automated recognition of melanoma using dermoscopic images. In parallel they used a self-advised SVM algorithm to enhance classification results by counteracting the effect of misclassified data. They tested the proposed model on a collection of 100 dermoscopic images and then compared the classification performance with some popular classification methods and the proposed model using the deep neural processing outperforms most of the popular techniques including KNN, ANN, SVM and semi supervised algorithms like Expectation maximization and transductive SVM. After going through the experimental results they computed the accuracy as 89%.

In Paper [10] authors proposed an automated system for detecting melanoma skin cancer from plain photographs of affected skin regions. They applied the ABCDEs rule for detecting melanoma as it is used in most of the cases. They used the Grab Cut algorithm to segment an input image into lesions of interest appeared to be melanoma. Then they extracted some features such as the shape, color, and geometry by using image processing techniques. They used Gaussian radial basis kernel (SVM·RBF) to categorize the extracted features as cancerous "malignant" or non-cancerous mole "benign" .It was found from the results that only six features can be sufficient to detect melanoma. The accuracy after the statistical analysis was computed as 86.67% .

In paper [11] authors proposed asystem, which focus on the problem of skin lesion classification. The proposed solution is built around the VGGNet convolutional neural network architecture and uses the transfer-learning paradigm. Experimental results are as follows: on the ISIC Archive dataset, the proposed method achieves a sensitivity value of 78.66%.

III. TOOLS AND TECHNIQUES USED

Python

Python is widely used general-purpose, high level programming language. It was designed by Guido van Rossum in 1991 and developed by Python Software Foundation. The main purpose for it's development was for emphasis on code readability and its syntax allows programmers to express concepts that to in fewer lines of code. It is a programming language that lets you work quickly and integrate systems in a moreefficient manner.

Libraries in Python

- Numpy
- Keras
- Pandas
- Matplotlib

Neural Network

Neural networks are a set of algorithms, modeled loosely after the human brain, that are designed to recognize patterns. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, be it images, sound, text or time series, must be translated.

Neural networks help us cluster and classify. You can think of them as a clustering and classification layer on top of the data you store and manage. They help to group unlabeled data according to similarities among the example inputs, and they classify data when they have a labeled dataset to train on. (Neural networks can also extract features that are fed to other algorithms for clustering and classification; so you can think of deep neural networks as components of larger machine-learning applications involving algorithms for reinforcement learning, classification and regression.)

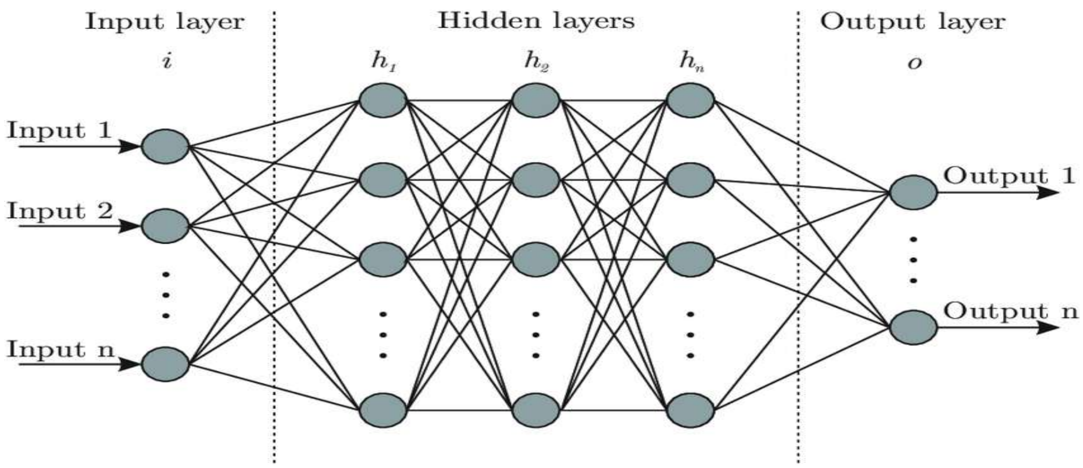


FIG 1: Structure of Neural Network

International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252

Convolutional Neural Network

A specific kind of such a deep neural network is the convolutional network, which is commonly referred to as CNN or ConvNet. It's a deep, feed-forward artificial neural network. Remember that feed-forward neural networks are also called multi-layer perceptrons(MLPs), which are the quintessential deep learning models. The models are called "feed-forward" because information flows right through the model. There are no feedback connections in which outputs of the model are fed back into itself.

CNNs specifically are inspired by the biological visual cortex. The cortex has small regions of cells that are sensitive to the specific areas of the visual field. This idea was expanded by a captivating experiment done by Hubel and Wiesel in 1962. In this experiment, the researchers showed that some individual neurons in the brain activated or fired only in the presence of edges of a particular orientation like vertical or horizontal edges. For example, some neurons fired when exposed to vertical sides and some when shown a horizontal edge. Hubel and Wiesel found that all of these neurons were well ordered in a columnar fashion and that together they were able to produce visual perception. This idea of specialized components inside of a system having specific tasks is one that machines use as well and one that you can also find back in CNNs.

Convolutional neural networks have been one of the most influential innovations in the field of computer vision. They have performed a lot better than traditional computer vision and have produced state-of-the-art results. These neural networks have proven to be successful in many different real-life case studies and applications, like:

- Image classification, object detection, segmentation, face recognition.
- Self driving cars that leverage CNN based vision systems.

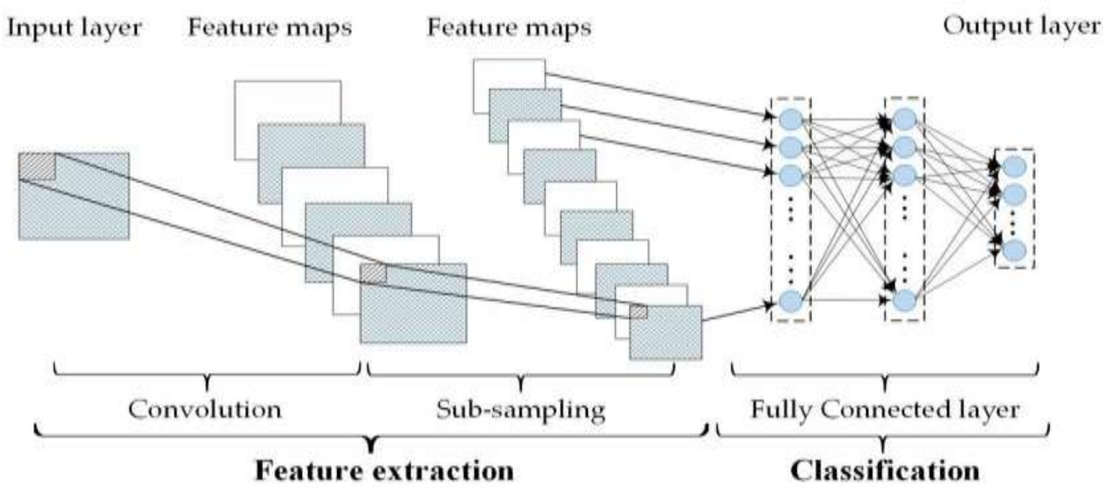


FIG 2: Convolutional Neural Network

VGGNet 16

VGG16 is a convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition". It was one of the famous models submitted to ILSVRC-2014. The model achieves 92.7% top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes.

This network is characterized by its simplicity, using only 3×3 convolutional layers stacked on top of each other in increasing depth. Reducing volume size is handled by max pooling. Two fully-connected layers, each with 4,096 nodes are then followed by a softmax classifier (above).VGG16 was trained for weeks and was using NVIDIA Titan Black GPU's.

The architecture depicted below is VGG16.

International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252

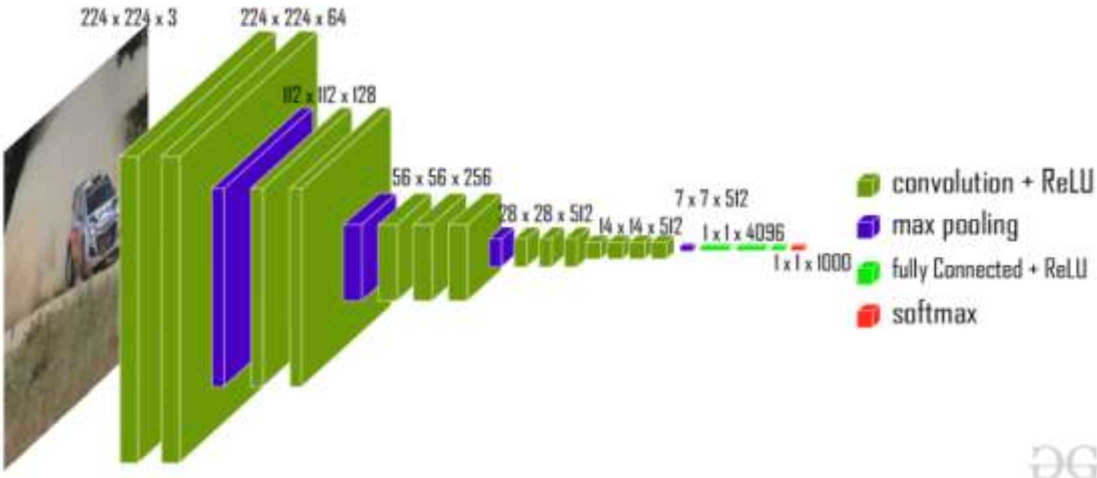


FIG 3 : VGGNet 16

ResNet 50

At last, at the ILSVRC 2015, the so-called Residual Neural Network (ResNet) by Kaiming He et al introduced novel architecture with "skip connections" and features heavy batch normalization. Such skip connections are also known as gated units or gated recurrent units and have a strong similarity to recent successful elements applied in RNNs. Thanks to this technique they were able to train a NN with 152 layers while still having lower complexity than VGGNet. It achieves a top-5 error rate of 3.57% which beats human-level performance on this dataset.

Even though ResNet is much deeper than VGG16 and VGG19, the model size is actually substantially smaller due to the usage of global average pooling rather than fully-connected layers — this reduces the model size down to 102MB for ResNet50.

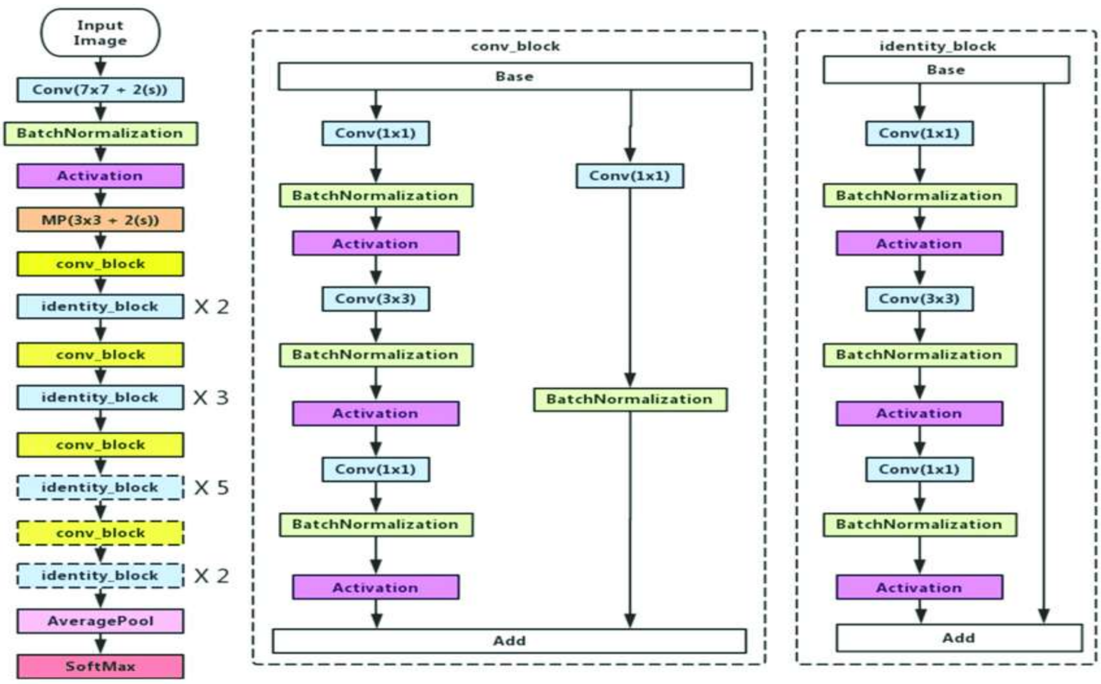


FIG 4: ResNet 50

International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252

IV. METHODOLOGY

Dataset description

The dataset consists of 1535 dermatoscopic images which are released as a training set for academic machine learning purposes and are publicly available through the ISIC archive. This benchmark dataset can be used for machine learning and for comparisons with human experts. They collected dermatoscopic images from different populations and stored by different modalities. Given this diversity we had to apply different cleaning methods and developed semi-automatic specifically trained neural network for its cleaning. Cases include a representative collection of all important diagnostic categories in the realm of pigmented lesions. More than 50% of lesions have been confirmed by pathology, while the ground truth for the rest of the cases was either follow-up, expert consensus, or confirmation by in-vivo confocal microscopy.

The specific dataset to use is:
ISIC_MSK-2: Benign and malignant skin lesions. Biopsy-confirmed melanocytic and non-melanocytic lesions.

- Benign:1167
- Malignant:352
- Indeterminate : 9 (Not used)
- Indeterminate/benign : 1 (Not used)
- Indeterminate/malignant : 2 (Not used)
- Nan : 4 (Not used)

Total number of images in dataset 1535
Dimensions of image 1024 X 768
As summary the total images to use are:

Benign Images	Malignant Images
1167	352

Number of images used for training the classifier:1215 (80%)
Number of images used for testing the classifier: 304 (20%)
Some sample images are shown below:
1.Sample image of benign moles:

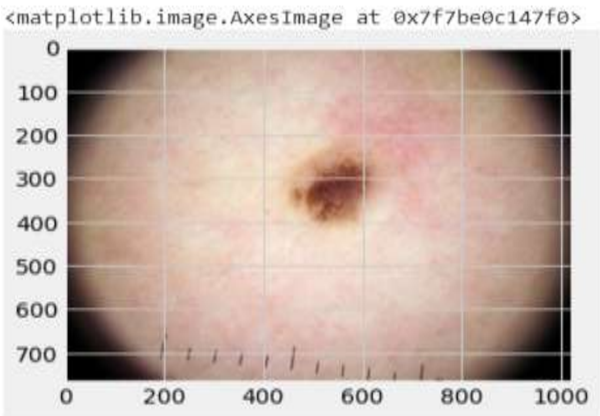
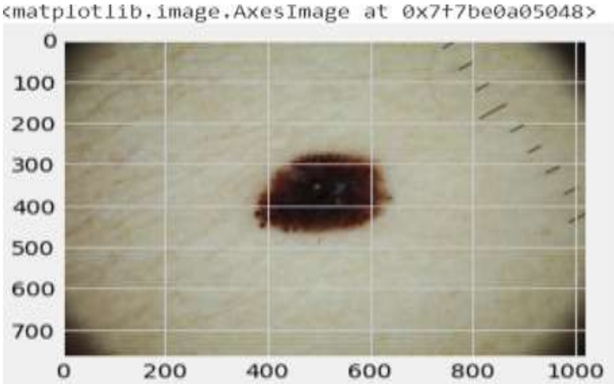


Fig 5: Sample image of malignant moles

International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252

<matplotlib.image.AxesImage at 0x7f7be0a05048>



General Architecture of Model

The Proposed model helps the dermatologist by considering the image of mole, can calculate the probability that a mole can be malign.

In this we created a Convolutional Neural Networks (CNNs) model which automates most of the diagnosis process with equal or more accuracy than the current methods.In this model we replicated a CNN using 1535 training images to compare the results to human experts and analyze the results to see how they contrast.

In upcoming steps we will show how we built this model, as well as the final results and comparison among various models.

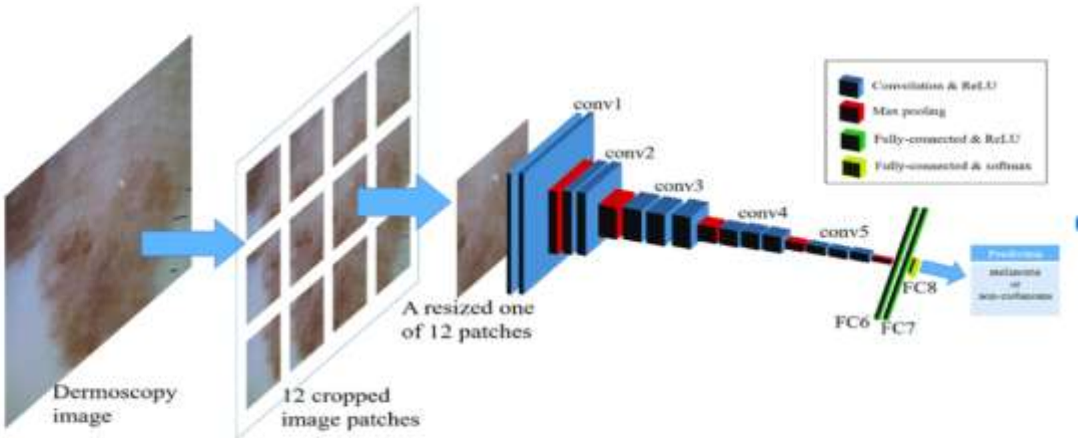


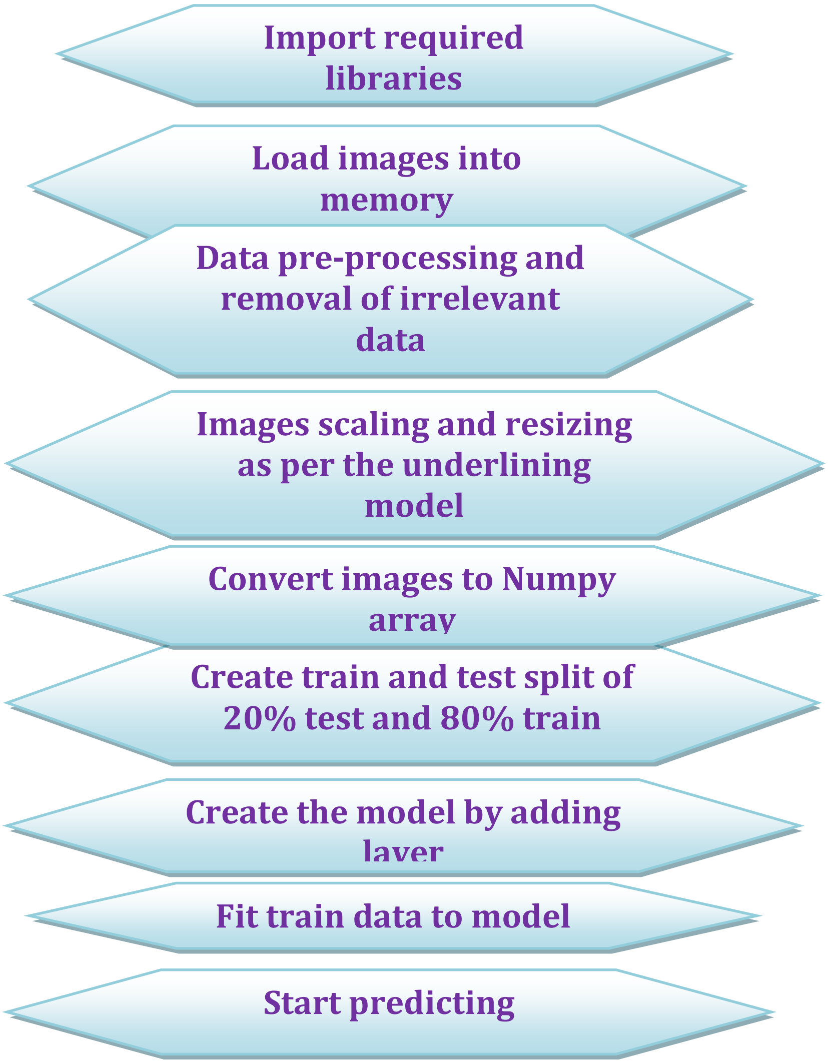
Fig 6: Schematic overview of CNN architecture

The dataset we used includes 2 major categories of skin lesions, benign and malignant lesions. Proposed model uses the ISIC_MSK-2dataset and is processed on a GPU.

Proposed Approach

Figure 4.4 depicts the flow of program and the different steps which are involved in model creation and prediction along with the different preprocessing methods.

International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252



Importing Essential Libraries
These libraries include Matplotlib, Numpy, Pandas, Sklearn and Keras.

Reading Data

```
metadata = pd.read_csv('metadata.csv')
print(metadata.shape)
```

```
metadata.sample(10)
```

	_id	name	meta.clinical.age_approx	meta.clinical.anatom_site_general	meta.clinical.benign_malignant	meta.clinical.clin_size_long_diam_mm
577	[illegible]	[illegible]	5	upper extremity	benign	5.5
1408	[illegible]	[illegible]	35	head/neck	malignant	6.0
782	[illegible]	[illegible]	35	upper extremity	benign	2.0
681	[illegible]	[illegible]	40	upper extremity	benign	5.0
96	[illegible]	[illegible]	15	lower extremity	benign	8.0
986	[illegible]	[illegible]	30	anterior torso	malignant	5.0
866	[illegible]	[illegible]	55	lower extremity	malignant	3.0
575	[illegible]	[illegible]	55	anterior torso	benign	3.0
473	[illegible]	[illegible]	40	anterior torso	benign	5.0
573	[illegible]	[illegible]	65	head/neck	malignant	4.0

Preprocessing

The invalid and null values from a csv file containing the labels for the respective images are removed and data is cleaned. Invalid value in this context was a label 'indeterminate', nan, 'indeterminate/benign', 'indeterminate/malignant'marked in the csv. After loading images in the memory, the different possible transformations are rotation, width shift, height shift, rescale, shearing, zoom and horizontal fit in order to create an image from different perspectives. In the proposed system only scaling of the image is used because further transformations reshaped the image to a perspective which is impossible for an input image.

Exploratory data analysis (EDA)

This step is to better visualize and understand the data for ourselves- seeing the different features of the dataset, how it is distributed, and some actual numbers.

Here we saw the distribution of data by category.

Label Encoding

The loaded images are converted into a numpy array and the respective csv file is labeled into two different classes, 0 for benign and1 for malignant.

Loading & Resizing of images

The images are loaded into the img column and resized.

Train Test split

The data is split into training and testing set with 80 20 division respectively.

Model Building

Here comes the fun part, using Keras sequential API and adding one layer at a time staring with input. The part that really separates CNN's from other Neural Nets.

A quick reminder of the 4 steps of CNNs:

Convolution -> pooling -> flatten -> full connection

Basically….

- Start off with aninput image.
- Apply filters or feature maps to the image, which gives a convolutional layer.
- Breakup the linearity of the image using rectifier function
- The image is ready for the pooling step.
- After pooling is done, it ends up with a pooled feature map.
- The pooled feature map is flattened before inserting it into an artificial neural network.

Setting Optimizer and loss function

The model is compiled with an 'adam' optimizer and a 'binary_crossentropy' loss function(error rate between the observed labels and the predicted ones) – helps to measure how poorly our model performs on images with known labels.

Fitting the model

The data separated for training and testing is used in fitting and training of the model. The testing data is utilized in validation of model. It is trained using the default batch size and an epoch value of 40. The accuracy and loss are calculated using the following functions.

$$Accuracy = \frac{truepositive + truenegatives}{totalexamples} \quad (1)$$

$$L(y,\hat{y}) = -\sum_{j=0}^{M}\sum_{i=0}^{N}(y_{ij} * log(\hat{y}_{ij})) \quad (2)$$

Further enhancements

In order to improve our model performance and compare it with the existing image classification techniques, transfer learning[14] is used and the convolution layers of different models[14] like VGGNet16, VGGNet19, RESNet50, Xception and InceptionRESNetV2 pre trained on imagenet dataset[15] are utilized. The convolution layers of the selected layers are frozen and the dense layer of the pre trained models is replaced with a custom dense layer and trained with the available ISIC_MSK-2dataset[B].

V. RESULTS AND OBSERVATIONS

International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252

A dataset of 1535 images of size 1024x768 is taken and split into training set of 1215 images and a validation set of 304 images. The images are resized according to the underlying model into three categories i.e. 128x128, 256x256 and 512x512. A CNN model is prepared and trained with the training data and the results are compared with the available image classification models, trained using transfer learning technique, on the same data. The number of epochs for each model (chosen based on examining the behavior of accuracy/loss plots vs. number of epochs) was: 20 epochs for CNN, 10 epochs for VGG16 and RESNet50. Validation is done using 304 validation image data.

There are 3 models, which are taken to compare their results on different resolutions. Table 5.1 and Table 5.2 show the accuracy and the loss comparison of all the models that are trained on the ISIC_MSK-2 dataset.

The model evaluation is performed on the training and testing partition of the ISIC_MSK-2 dataset.

Comparison of results on training dataset

After comparing results,we can say that highest training accuracy for resolution 256x256 is obtained by RESNet50 which is 0.9613

S. No.	Model Used	Resolution	Loss	Accuracy
1.	CNN	128 x 128	0.4923	0.7770
2.	CNN	256 x 256	0.4857	0.7901
3.	CNN	512 x 512	0.5381	0.7712
4.	VGGNet16	128 x 128	0.2715	0.8889
5.	VGGNet16	256 x 256	0.1903	0.9267
6.	VGGNet16	512 x 512	0.1546	0.9432
7.	RESNet50	128 x 128	0.2743	0.9012
8.	RESNet50	256 x 256	0.1890	0.9613

Table 1: Model Evaluation: training dataset

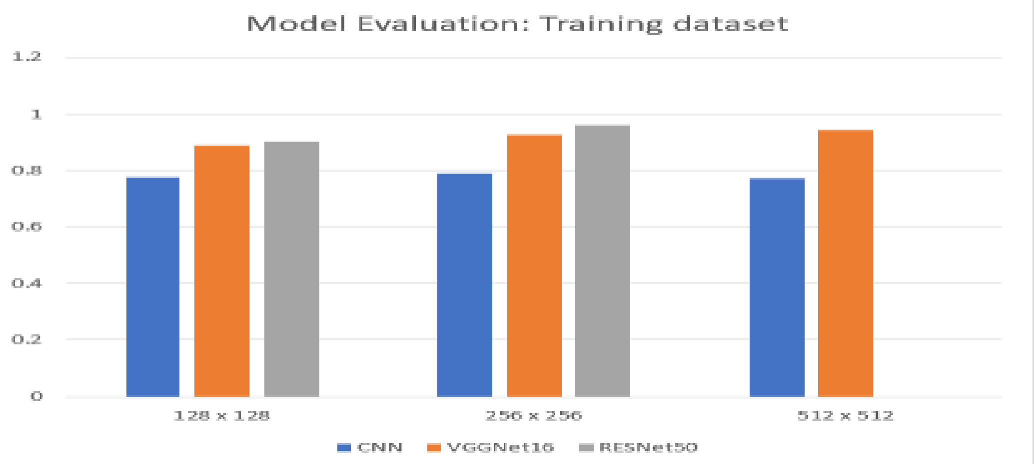


Fig 7:Training Accuracy comparison of CNN, RESNet50 and VGGNet16

Comparison of results on validation dataset

After comparing results, we can say that highest training accuracy for resolution 128x128 is obtained by VGG16, which is 0.8191

S. No.	Model Used	Resolution	Loss	Accuracy
1.	CNN	128 x 128	0.4690	0.7928
2.	CNN	256 x 256	0.5365	0.7895
3.	CNN	512 x 512	0.5553	0.7566

International journal of advances in engineering and management (IJAEM)
Volume 3, issue 6 June 2021, pp: 1837-1848 www.ijaem.net ISSN: 2395-5252

4.	VGGNet16	128 x 128	0.4716	0.8191
5.	VGGNet16	256 x 256	0.5122	0.7928
6.	VGGNet16	512 x 512	0.6067	0.7664
7.	RESNet50	128 x 128	1.1297	0.7829
8.	RESNet50	256 x 256	0.5478	0.7632

Table 2: Model Evaluation: Validation dataset

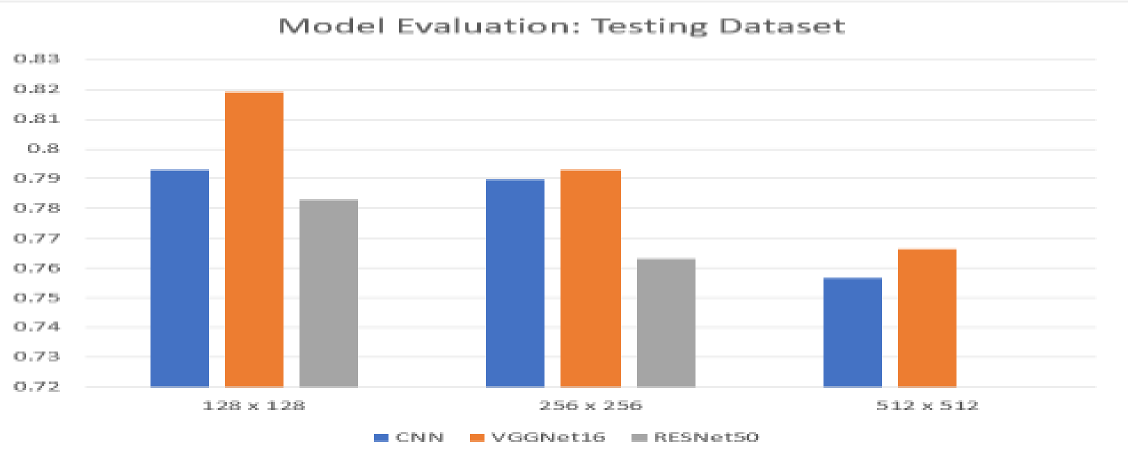


Fig 8: Validation Accuracy comparison of CNN, RESNet50 and VGGNet16

VI. CONCLUSION AND FUTURE SCOPE

In this research work, prediction of skin lesion image is performed with the aim to help the patient to identify the skin cancer without going to hospitals. This diagnosis research work uses different types of neural networks to predict skin cancer.

In this work, we used RESNet50, CNN and VGGNet16 neural networks for prediction of skin cancer and with VGGNet16 we have achieved highest accuracy of 81.91% .

Its accuracy can be increased by tuning or by adding some new samples to the dataset. We can also do experiments with different pooling, architectures and optimizers because these things can make a big difference in model behavior

As a future work, data augmentation techniques can be used to tackle the unbalanced classes issue.

REFERENCES

[1]. N. C. Lynn and Z. M. Kyu, “Segmentation and Classification of Skin Cancer Melanoma from Skin Lesion Images,” 2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), 2017.

[2]. S. Mustafa, A. B. Dauda, and M. Dauda, “Image processing and SVM classification for melanoma detection,” 2017 International Conference on Computing Networking and Informatics (ICCNI), 2017.

[3]. M. E. Celebi, H. A. Kingravi, Y. A. Aslandogan, and W. V. Stoecker, “Detection of blue-white veil areas in dermoscopy images using machine learning techniques,” Medical Imaging 2006: Image Processing, Feb. 2006.

[4]. I. A. Ozkan and M. Koklu, “Skin Lesion Classification using Machine Learning Algorithms,” International Journal of Intelligent Systems and Applications in Engineering, vol. 4, no. 5, pp. 285–289, 2017.

[5]. Z. Waheed, A. Waheed, M. Zafar, and F. Riaz, “An efficient machine learning approach for the detection of melanoma using dermoscopic images,” 2017 International Conference on Communication, Computing and Digital Systems (C-CODE), 2017.

[6]. Shalu and A. Kamboj, “A Color-Based Approach for Melanoma Skin Cancer Detection,” 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), 2018.

[7]. A. Masood and A. Al-Jumaily, “Semi-advised learning model for skin cancer diagnosis based on histopathalogical images,” 2016 38th Annual International Conference of the IEEE Engineering in

Medicine and Biology Society (EMBC), 2016.

[8]. R. S. S. Sundar and M. Vadivel, “Performance analysis of melanoma early detection using skin lession classification system,” 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), 2016.

[9]. A. Masood, A. A.-Jumaily, and K. Anam, “Self-supervised learning model for skin cancer diagnosis,” 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER), 2015.

[10]. S. Mustafa and A. Kimura, “A SVM-based diagnosis of melanoma using only useful image features,” 2018 International Workshop on Advanced Image Technology (IWAIT), 2018.

[11]. Lopez, Adria Romero, et al. "Skin lesion classification from dermoscopic images using deep learning techniques." 2017 13th IASTED international conference on biomedical engineering (BioMed). IEEE, 2017.

[12]. www.medium.com

Data Citation

1. https://www.isic-archive.com/#!/topWithHeader/onlyHeaderTop/gallery