* **Aakash Shanavas**
* **Riya Bansode**
* **Prajakta Ingle**
* **Varsha Shashidhar**

SKIN CANCER – MALIGNANT OR BENIGN DETECTION

Contents

[Introduction 2](#_Toc87898636)

[Methodology 2](#_Toc87898637)

[Proposed Method 3](#_Toc87898638)

[Process Flow 3](#_Toc87898639)

[Preprocessing 3](#_Toc87898640)

[Save the preprocessed file 3](#_Toc87898641)

[Feeding the preprocessed data to convolutional neural network (CNN) 3](#_Toc87898642)

[• Convolution layer 4](#_Toc87898643)

[• Pooling layer 4](#_Toc87898644)

[• Fully connected layer 4](#_Toc87898645)

[Train 4](#_Toc87898646)

[Save the 4](#_Toc87898647)

[Prediction 4](#_Toc87898648)

[Conclusion 10](#_Toc87898649)

[Future Scope: 10](#_Toc87898650)

[References: 11](#_Toc87898651)

Introduction**:**

Cancer has adversely impacted physical and mental health of people all over the world. In 2020 it was estimated that 1,806,590 new cases of cancer will be diagnosed in United States alone leading to an astounding 606,520 deaths. Five most prominent cancers seen in patients based on the estimated new cases in 2020 are breast cancer, lung and bronchus cancer, prostate cancer, colon and rectum cancer and melanoma of the skin. [2]

At some point in our education, we have studied human anatomy and the different organs that make up our human body. Every organ in our body be it brain, skin, liver, kidney, bladder, heart, or intestine comprises of different types of cells. Cells have a peculiar ability to reproduce themselves by dividing which helps our body grow. The mechanism that enables and regulates this division is still arcane in the medical world. In some individuals there are formation of cells which grow uncontrollable and do not build typical bone tissues or flesh, which the individual’s body is used to. These cells divide swiftly in a disorderly manner which piles up into an unstructured mass or tumor. Some tumors remain in the part where they begin to grow, not spreading to other parts of the body. These are known as benign tumors. There are some others which not only destroy the part of the body where they originate but also cause new growth to other parts effecting vital organs which most likely kills the patient. These are known as malignant tumors. [3]

Skin cancer being one of the most common forms of cancer, results in enormous expenditure in United states itself. An examination of Medical Expenditure Panel Survey Data shows that every year 4.3 million adults are treated for the most common types of skin cancer including basal cell and squamous cell carcinomas which amounts to an astounding $4.8 billion. [4]

Since skin cancer takes up such a huge chunk of expenditure of any country’s economy, health organization including WHO have ramped up their efforts alongside research institutes and top medical field practitioners to develop various methods and techniques to help better and early diagnosis in patients. Models are being developed using Machine learning algorithms including neural networks which are put that help tremendously in these efforts. [5-8]

A paper published in scientific journal called Nature shows deep neural network models being able achieve accuracy of board-certified dermatologist in classification of skin cancer. Mobile devices including phones, tablets etc., outfitted with deep neural networks can help diagnosis being made available and accessible outside clinics to the remotest of the areas. Projection shows nearly 6.3 billion smartphone subscriptions by the year 2021 which can potentially provide relatively inexpensive access to crucial diagnostic care. [9]

Through this project our focus is on developing Machine learning models that can accurately predict whether the image tested is Benign or malignant form of skin cancer. The dataset used for training our models contains a balanced number of images of benign skin moles and malignant skin moles. We are using different architectures of CNN algorithms to build models which are subsequently trained and tested using the images in the dataset.

# Methodology

To determine whether a patient has skin cancer, he or she must have a single examination by a dermatologist. This platform enables dermatologists to handle a wide range of situations much more quickly than they would otherwise. There have been a variety of symptomatic checklists developed. One of the checklists is ABCDE, which includes items such as - Asymmetry(A) - One half of the afflicted cell that has become a tumor does not coordinate with the other half. Border(B)- The tainted cells' edges/fringe get damaged, scored, and concealed. Color(C): The shade isn't consistent. There are tan or dark colored blotches on the skin, as well as dark. The ugly aspect is enhanced with splashes of red, white, and blue. Diameter(D)- The cell width is greater than or equal to 6mm. Evolution(E): Malignant Melanoma has evolved because of the previously described modifications or improvements. [14]

# Proposed Method

This method employed the tagged images "benign" and "malignant." There are three layers in our proposed system. The input layer is the first layer, which is where the data sets are trained. The input layer collects data and adds some weight to it, which is then passed on to hidden levels. To detect a pattern, the neurons in the hidden layer separate the characteristics from the data. The pattern is then used to create output layers that select appropriate classes based on the pattern. Finally, binary classification is utilized to identify appropriate classes 1 and 0. In our situation, class 0 denotes the absence of hazardous cells, whereas class 1 denotes the presence of malignant cancerous cells. [14]

Diagram, engineering drawing

Description automatically generated

[13] Page 3

# Process Flow

## Preprocessing

The data set comprising photos could span exceptionally huge in width and height. The picture's width is 1022 pixels, and its height is 767 pixels, making it extraordinarily large to analyze and necessitating significantly greater computer capacity to register many images, which takes a long time and wastes memory. We must reduce the size of the information pictures so that our machine can process them with less memory and graphical computational resources. Another issue is with the size of the image. If the image size is 1024\*1024\*3, the feature size for calculation to transmit it to a deep neural network, particularly a convolutional neural network, will be enormous. It will be defined in such a way that just one-color channel remains to address these two issues while interpreting the photos.[14]

## Save the preprocessed file

Each of the preprocessed photos, as well as their classes, are preserved in the record. For further processing, benign and malignant photos are selected from the dataset. The photos that do not have a class label are discarded. Finally, the captured images are fed into a convolutional Neural Network for processing.[14]

## Feeding the preprocessed data to convolutional neural network (CNN)

In a convolutional Neural Network, there are three sorts of layers. That can be found in the following section:

## • Convolution layer

A Convolutional Neural Network's main building piece is the CONV layer. The CONV layer parameters are made up of a set of kernels which are nothing but K filters, each of which has a width and height and is almost always square. The depth of a CNN input is the number of channels in the image (which is three when working with RGB). The number of filters used in the previous layer will determine the depth of volumes deeper in the network. The network "learns" filters that activate when they see a specific sort of feature at a specified spatial point in the input volume in this manner. When filters in lower layers of the network see edge-like or corner-like regions, they may activate. When using CNNs, we select to link each neuron to only a small portion of the input volume, which we refer to as the receptive field. After each CONV layer in a CNN, we apply a nonlinear activation function, such as ReLU. Activation layers aren't technically "layers," but they're often overlooked in network architecture diagrams because it's thought that an activation follows a convolution quickly. [15]

## • Pooling layer

## The POOL layer's principal role is to gradually shrink the input volume's spatial dimension. POOL layers use the max or average function to operate on each of the depth slices of an input separately. To reduce spatial size, max pooling is usually done in the middle of the CNN architecture, whereas average pooling is usually done as the network's last layer. [15]

## • Fully connected layer

The features obtained from pooling layers are flatten to a single layer so that we can feed the model to a fully connected neural network. Neurons in fully connected layers are fully linked to all activations in the previous layer. These layers are always at the network's end.[15]

## Train

The final step is to train. Our model will need to be trained 200 times. The system's loss decreases to a specific amount at regular intervals.[15]

Save themodel

The model is kept for further testing. After that, the model is used to estimate whether photos are likely to include malignant or benign images.[15]

## Prediction

Using the final output layer, we must anticipate the images. We evaluate our system using the accuracy on training and testing data set, ROC curve, Sensitivity and Specificity measures after we forecast the testing images.[15]

System Performance and Results:

|  |  |  |  |
| --- | --- | --- | --- |
| Sr.No. | Model Name | Accuracy on training set | Accuracy on testing set |
| 1 | Deep Learning model using TensorFlow Hub with Keras with 4 epochs | 83.17% | 80.5% |
| 2 | Deep Learning model using TensorFlow Hub with Keras with 30 epochs | 88.66% | 76% |
| 3 | CNN model using ResNet50 | 99.67% | 54.54% |

In this study, 2137 training images and 500 validation images were used in the training model. The images obtained from the ISIC dataset, consist of two classes benign or malignant. These images were trained using the Deep Learning model using TensorFlow Hub with Keras and CNN model using ResNet50 with various optimizer methods such as RMSprop and Adam optimizer with a learning rate of 0.001 and used loss categorical cross-entropy. The performance parameters measured in this study are accuracy, specificity, and sensitivity.

Iterations:

* 1. Model 1: CNN model using ResNet50

Initially we used CNN model with ResNet50 which is a ResNet50 is a variant of [ResNet model](https://iq.opengenus.org/resnet/) which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. In general, in a deep convolutional neural network, several layers are stacked and are trained. The network learns several low/mid/high level features at the end of its layers. In residual learning, instead of trying to learn some features, we try to learn some residual. Residual can be simply understood as subtraction of feature learned from input of that layer. ResNet does this using shortcut connections (directly connecting input of nth layer to some (n+x)th layer. It has proved that training this form of networks is easier than training simple deep convolutional neural networks and the problem of degrading accuracy is resolved [[12]](https://iq.opengenus.org/resnet/). Using CNN models with ResNet50 gives a better accuracy on training set , but the model failed to achieve a better accuracy on test data set. This led the model to predict only one class.

In our learning, we tried to gradually increase the data sample set for training and test and then evaluating the model, but it gave predictions on only one class. In future, we aim to get more data images of the other class as well and train the model to get better test accuracy.

For this reason we decided to use our second model which explains as follows.

* 1. Model 2: Deep Learning model using TensorFlow Hub with Keras with 4 epochs:

In this study, 2000 training images and 150 validation images were used in the training model. We used model with adding the Tensorflow Hub Library with Keras and load the InceptionV3 architecture along with its [ImageNet](http://www.image-net.org/) pre-trained weights. With the optimizer as RMSprop and activation function as sigmoid, we trained the model with 4 epochs which could achieve an accuracy of 83.17% on training data set and 78.12% on validation data set.

Accuracy on Training set:

Graphical user interface, text

Description automatically generated

Accuracy on Test set:

Graphical user interface, text, application

Description automatically generated

From the confusion matrix that we obtained, we can infer that

Chart, treemap chart

Description automatically generated

1. TP (True Positive): Percentage of patients who have been properly classified to have benign tissues, meaning they do not have harmful cancerous cells is 82%.
2. TN (True Negative): Percentage of patients who are correctly classified with malignant tissues, meaning they have cancerous cells and have high risk the disease is 49%.
3. FP (False Positive): Percentage of predicted patients who do not possess harmful malignant cells disease, but they are suffering with malignant cancerous cells is 51%.
4. FN (False Negative): Percentage of patients predicted as they have malignant cancer cells but, they are not suffering from the disease is 18%.

Sensitivity: So, our model gets about 49% probability of a positive test given that the patient has the disease (bottom right of the confusion matrix), that's often called sensitivity. So in our example, out of all patients that have a malignant skin disease, we successfully predicted 49% of them as malignant. This needs improvement as our model needs to be trained precisely.

Specificity : The other metric is specificity, you can read it in the top left of the confusion matrix, we got about 82%. In our project, out of all patients that has a benign, we predicted 82% of them as benign. With high specificity, the test rarely gives positive results in healthy patients, whereas a high sensitivity means that the model is reliable when its result is negative.

Text, letter

Description automatically generated

ROC curve:

Graphical user interface

Description automatically generated with medium confidence

Images misclassified by the model:

Calendar

Description automatically generated with low confidence

Images classified correctly by the model:

A picture containing calendar

Description automatically generated

* 1. Model 3: Deep Learning model using TensorFlow Hub with Keras with 30 epochs:

As we learnt that training the model for more epochs would help the model to better understand the malignant cases and thus reduce the risk of misclassification for wrong predictions, we also executed the model with 30 epochs and we were able to achieve better results as follows:

Accuracy on training set:

Text, application

Description automatically generated with medium confidence

Diagram

Description automatically generated with medium confidence

Accuracy on testing set :

Graphical user interface, text

Description automatically generated

Confusion Matrix:

Chart, treemap chart

Description automatically generated

Here we can see an improvement in the prediction of Malignant cases and decrease the percentage of misclassifying true malignant as benign. Thus, more model will get trained we can get better results.

Images misclassified by the model:

A picture containing calendar

Description automatically generated

Images correctly classified by the model:

Calendar

Description automatically generated

Conclusion:

A cancer patient goes through a lot, and we must support them in their time of need. Early detection of the type of cancer ensures that it can be treated, if not cured. The CNN is more closely tied to computer vision than other types of neural networks, it performs better when identifying picture data than ANNs, CNNs, KNNs, and RBFNs for classification of lesion images. The majority of skin cancer detection research focuses on determining if a particular lesion image is malignant.

This model is proposed with Deep Learning model using TensorFlow Hub with Keras with **RMSProp** optimizer provides the best performance in classifying the dataset of skin cancer lesions and benign skin lesions with **88.66**% accuracy on training and **76%** accuracy on testing. The system reveals that the provided model is promising to use as an existing tool for medical staff in determining the diagnosis of skin cancer is benign or malignant, based on the performance outcomes. In further research, systems can be developed to classify the various types of skin cancer and other skin diseases, which we will discuss in the future scope below.

# Future Scope:

According to certain research, CNN algorithms have already surpassed dermatologists' classification efficacy, and AI classifiers may soon attain sufficient sensitivity and specificity to shoulder the screening burden for detecting malignant skin malignancies. As a result, while some physicians may see AI as a danger, we believe it is nothing more than a diagnostic aid system, owing to the numerous limitations described in prior studies and the difficulty in making performance comparisons within published results.

Below are some possible ways that can be worked upon in future:

* When we are considering the model for predicting the disease as malignant or benign, we need a way to predict even more malignant cases even that we have very few malignant samples compared to benign. We can determine a good threshold to classify the level of intensity of the disease in the CNN model. Even can be predicted using different optimizers and algorithms.
* In the real world, it is rare to train a CNN from scratch, as it is hard to collect a massive dataset to get better performance. Instead, it is common to use a pre-trained network on a very large dataset and tune it for your classification problem, this process is called Transfer Learning. [11]
* The model can use it on other skin disease classifications (such as melanocytic classification), so we can add more skin diseases and use it for other problems as well.
* Transfer learning is a machine learning method where a model is trained on a task that can be trained (or tuned) for another task such as
* Here are the most important benefits of transfer learning: Speeds up training time, it requires fewer data, Use the state-of-the-art models that are developed by deep learning experts.
* Aside from novelty detection, there are several enhancements that can be done like Color constancy and lesion segmentation, for example, can be shown to aid in dermoscopic picture classification. It's also possible that incorporating lesion meta-data will improve diagnosis accuracy. Layer activations, t-SNE, and saliency maps are examples of techniques for comprehending and visualizing CNNs that could aid in the interpretation of models and the identification of features that discriminate between different types of lesions leading to skin cancer.
* The lightweight CNN models can be employed in on mobile or edge devices for the firsthand screen of the lesions to determine the if it is a cancerous lesion and with what pattern, and the intensity of it.

# References:

[1] <https://www.canaryfoundation.org/2012/08/13/how-long-have-we-been-plagued-by-cancer/>

[2] <https://www.cancer.gov/about-cancer/understanding/statistics>

[3] <https://seer.cancer.gov/archive/training/manuals/Book2.pdf> p 5.

[4] <https://www.cdc.gov/cancer/skin/statistics/index.htm>

[5]<https://www.nature.com/articles/nature21056.epdf?author_access_token=8oxIcYWf5UNrNpHsUHd2StRgN0jAjWel9jnR3ZoTv0NXpMHRAJy8Qn10ys2O4tuPakXos4UhQAFZ750CsBNMMsISFHIKinKDMKjShCpHIlYPYUHhNzkn6pSnOCt0Ftf6> p 1.

[6] American Cancer Society. Cancer facts & figures 2016. Atlanta, American Cancer Society 2016. <http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-047079.pdf>

[7] Rogers, H. W. et al. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. JAMA Dermatology 151.10, 1081–1086 (2015)

[8] Stern, R. S. Prevalence of a history of skin cancer in 2007: results of an incidence-based model. Arch. Dermatol. 146, 279–282 (2010).

[9] <https://cs.stanford.edu/people/esteva/nature/>

[10] <https://www.mdpi.com/1660-4601/18/10/5479/pdf>

[11] <https://www.thepythoncode.com/article/skin-cancer-detection-using-tensorflow-in-python>

[12] <https://iq.opengenus.org/resnet/>

[13] Convolutional Neural Network (CNN) for Automatic Skin Cancer Classification System - Yunendah Nur Fu’adah et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 982 012005

[14] Skin Cancer Detection Using Convolutional Neural Network

[(PDF) Skin Cancer Detection Using Convolutional Neural Network (researchgate.net)](https://www.researchgate.net/publication/334751850_Skin_Cancer_Detection_Using_Convolutional_Neural_Network)

[15] [Convolutional Neural Network. Learn Convolutional Neural Network from… | by dshahid380 | Towards Data Science](https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529)