## Skin Cancer Prediction Using Deep Learning

## Introduction

Because skin cancer is so deadly, it is considered to be among the most lethal types of cancer. Deoxyribonucleic acid (DNA) in cells is the cause of most cases of skin cancer, according to the National Cancer Institute. Skin cancer spreads across many other areas of the body, making it much more treatable in the early stages. Therefore, it is best diagnosed and cured in the early stages. This is especially critical due to the growing incidence of skin cancer, high death rate, and pricy medical care.

There are two types of skin cancer tumours: Malignant and Benign. Malignant skin cancer may develop quickly and spread across other areas of the body. Benign mole is not considered to be a high-risk kind of cancer since it develops, however, do not speared. With the normal sight of a human inspection, these moles cannot be seen precisely, and therefore inaccurate findings lead to inaccurate medication and possibly death. Early diagnosis of malignant skin cancer may lead to a higher survival percentage. Thus, for greater effectiveness and precision, automated detection is more dependable. It is estimated that 80% of all cancer diagnoses are due to skin cancer. Surviving rates are extremely high if cancer is detected early on and treated appropriately. As a result, finding out as soon as possible whether the patient's symptoms match up is critical. in the face of cancer or not. In the past, doctors have detected skin cancers solely by looking at the patient. However, because humans make mistakes, this often results in inaccurate detection. Even It's a difficult thing for doctors to say, especially when cancer is still in its early stages. This is what I mean. to automate the entire pipeline by using computer vision. It is possible to train a neural network using hundreds of photos from both benign and malignant categories and cancerous as well. The model can determine whether an image has non-linear interactions by learning about them. according to whether it's a benign and malignant form of the disease This automation may not only be more cost-effective, but it may also offer other advantages. more accurate results while also reducing the amount of time and manual labour spent Work that is beneficial to the organization is performed. Our system can be used in places where there aren't any qualified doctors. This project has the potential todrastic improvements in developing country human health care and well-being Africa.

In this project, we will try to classify malignant and benign cancer based on the images that are provided in the data. The data that is available to train the models consists of two subfolders: train, and test. Both the files consist of images that belong to both classes. In this project, we try to diagnose the skin cancer type with high accuracy.

## Business Problem

Skin Cancer being one of the deadliest diseases in the world that could end human life by effecting their skin. Every year thousands of people are being infected by this disease and it is highly important to identify this disease in the early stages to cure the patients. Because of the increase in the cases of Skin Cancer every year it is hard for limited number of health care officials to manually check and verify each patients. It would be very helpful if this process of manual inspection can be automated with high accuracy.

Technologies such as Deep learning could automate this process and identify the Skin Cancer in the early stages and could reduce the burden on the healthcare officials. Detecting the Skin Cancer can be possible by training an efficient Convolutional Neural Network model that could identify the Cancer efficiently.

## Dataset

The dataset that is being used is a secondary data set. This dataset is being collected from an open source platform called Kaggle. Kaggle is a platform for Data Scientists and Machine Learning users that provides huge amounts of data across different domains. This data can be used further for the purpose of research or for education.

This data was collected from patients. I have downloaded the dataset from Kaggle <https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign>. This dataset contains a balanced dataset of images of malignant skin moles and benign skin moles. The dataset contains two folders with each 1800 images(244\*244).

## Project Aim

In this project our aim is to detect the cancer using the convolutional neural networks by training these networks with different architectures and evaluate the networks based on the accuracy they have achieved. Our primary aim remains to build a classification model that could classify between the Cancer disease.

## 1.4. Project Objectives

1. Read the research papers on the healthcare that are closely related to this project and gain knowledge in the subject area.
2. Perform the CNN classification task on the dataset to classify the images based on the skin cancer.
3. Evaluate the performance of the CNN models that are based on different activation units.

## 1.5. Research Questions

Research questions are the important part of a study that allows the researchers to learn in-depth regarding a subject. The researchers aim to answer the research questions related to their work. The following research questions were identified in the project that will be analyzed during this project.

1. How efficient is the convolutional neural networks in healthcare industry?
2. Is the CNN efficient enough to diagnose the skin cancer based on the images?
3. What activation units are suitable for the CNN in this classification task?

In this project, these research questions play a key role in understanding the behaviour of the CNN models with different architectures, evaluate the performances of the models and interpret the outcomes at various architectures, and understand the strategies to improve the performances of the classifiers.

## 1.6. Techniques& Tools

The techniques that will be used to successfully complete this project are Python programming language, Deep Learning, Convolutional Neural Networks, Image-processing.

The tools that will be used in this project are:

Jupyter notebooks: The IDE in which we write the code

Anaconda: This is a distribution for Python programming.

Keras: Framework of the python that provides an interface for running the deep learning models in Python language.

NumPy: This is a library that allows the users to perform mathematical functions on the high dimensional data.

# 2. Legal, Ethical, and Social Issues

Legal Issues –

Since this algorithm that’s being trained on medical data and responsible for identifying skin cancer at an early stage the life of an individual is dependent on the predictions made by the algorithm. If the cancer is not diagnosed at early stages due to any fault in the algorithm may lead to the death of the individual suffering from cancer. This will be a legal problem since the life is lost due to the error made by the algorithm.

Ethical Issues –

Machine Ethics &Algorithmic Biases:

The bias behaviour in the model could result in less accurate predictions which cannot be used to diagnose the patients. Using an algorithm that is biased could misclassify most of the results and misguide the doctors resulting in improper treatment to the patients. This breaks the ethical laws.

Data Control & Monetization:

The data that is being collected for predicting skin cancer cannot be used for personal benefits or cannot be sold to third parties without the patients acknowledging it. Using the data for marketing the products of the other companies without the consent from the patients can be a violation of the ethical laws.

Social Issues –

With the outcomes from a biased model that may classify a healthy patient as infected could be a serious trouble. The patient might go through mental breakdown and could impact the person as well as the people around the person. This will harm the mental state of the person and their family members. This might also result in the discrimination in the society towards the person.

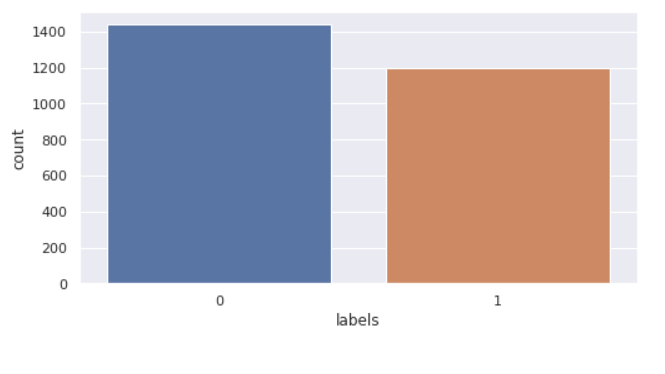
## 3. Methodology

A detailed explanation of the technique that was utilized to effectively accomplish this project is provided in this section of the report. I started by installing all of the necessary software, such as Python and Anaconda, and then I went on to install all of the libraries that were needed via the command prompt. NumPy, OS, CV2, Matplotlib, SciKit Learn, and Keras are some of the libraries that were used in the course of the research. The data was obtained via the Kaggle website. Thousands of picture files from chest X-rays are included inside the collection. There are two kinds of files in this collection: the Benign class and the Malignant class. In each set, there are files from the two classes that are present in the other files. All of the files are organized into two sets: Train and Test sets.

## 3.1. Importing Data

We are working on Jupyter environment since it has a very interactive interface, and it is suitable for production. This IDE is the most used by Data Scientists across the globe. The initial step was to import all the libraries in the notebook. Firstly, we have to import all the data that was collected from the Kaggle in order to proceed with our research. All the data will be imported to the Jupyter IDE.

I have used the OS library to read the directory of the data file location and I have imported the dataset and assigned them to train and test variables according to the way they were organized. Since the whole data was divided into two classes, I have plotted a bar plot to see how the data is distributed among the classes in the train data.



Here 0 indicates Benign class and 1 indicates Malignant class.

## 3.2. Image Pre-processing

After importing the data into the IDE, I have separated the labels and features and stored them in different variables for all the sets, i.e., train and test. We further perform the image pre-processing on the data to normalize and re-size the images. This step is intended to enhance the picture quality so where we can conduct a more thorough analysis of the image. The use of pre-processing allows us to eliminate unwanted distortions and improve certain characteristics that are essential for the specific application for which we are developing. Those characteristics may vary depending on the application.

There are two pre-processing stems that were performed as discussed. They are as follows:

### 

## 3.2.1. Normalization:

Data normalization is a crucial step in the image processing process since it guarantees that every input parameter (in this instance, pixel) has a comparable data distribution. As a result, the network's convergence time is reduced while it is being trained. Data normalization is accomplished by removing the average out of each pixel and dividing the resulting value by a standard deviation of the data set. In this case, the distribution of the data may resemble a Gaussian distribution with the origin at zero. We need positive pixel values for picture inputs, therefore we could choose to scale the normalized data in the range [0, 1] or [0, 255], depending on the application.

For the purpose of normalizing our data, I have transformed the features into a NumPy array and then divided the array by 255. In the train, validation, and test sets, this process was repeated for each feature set.

## 3.2.2. Resizing images

In computer vision, resizing pictures is a crucial step in the pre-processing process. Machine learning techniques learn to recognize patterns in smaller pictures, which is important since they train quicker on smaller images. The learning process for a twice-larger input images needs the network to learn over 4 times as much of pixels, which increases the amount of time it takes. A further point to consider is that many neural network model designs need input images to be of the same size, while our raw gathered images may be of different sizes.

It was necessary to scale the whole data set to fall inside the range of [-1, 1] across all of the feature sets. This ensures that the image's size falls within the boundaries of the specified range.

## 3.3. Data Augmentation

The quantity and variety of data provided during training have a significant impact on the accuracy of predictions made by Supervised Deep Learning models. When it comes to deep learning models and the quantity of training data needed, the relationship is similar to the relationship among deep learning models and the massive quantities of data required for the deep learning model to be successful. Computer vision tasks including such images categorization, object recognition, and segmentation have shown to be very effective among deep learning applications that are already in use today. In these types of applications, data augmentation may be utilized to successfully train deep learning models. Simple changes that may be done to an image include transformations like Flipping, Rotating, Translations, Cropping, and Scaling, as well as pixel transformations such as colour casting, varying contrast, and noise injection are a few examples.

The data augmentation helps the models to be robust since the model’s train on images that are transformed. The operations that were performed on the training data are as follows:

1. **Rotation range:** This operation ensures that the images are rotated at certain angles. The user specifies the angle for rotation.
2. **Zoom range:** This is an operation that transforms the images by performing the zoom in and zoom out.
3. **Width shift range:** It really causes the picture to be shifted horizontally towards left or right. If the value is variable and the range is more than one, the percentage of the entire width will be used as the range. Assume that the picture width is 100px. If width shift range = 1.0, the range will be from -100 percent to +100 percent, which is equivalent to -100px to +100px. It will move the picture in a random manner between these two ranges. A randomly chosen positive number will move the picture to the right side of the screen, while a randomly selected negative value would move the image towards the left of the screen. This may also be accomplished by choosing pixels. It will have the same impact if we specify width shift range = 100, which is the default value. More significantly, integer values more than one count pixel are used as range, whereas float values less than one count percentage of entire width are used as range.
4. **Height shift range:** It functions in the same way as width shift range but shifts vertically up or down.
5. **Horizontal flip:** This operation performs the random horizontal flip on the images to generate new images.

These whole operations were performed using the Keras image data generator function that allows to transform the images.

## 

## 3.4. Deep Learning - CNN

Artificial Intelligence has seen tremendous progress in recent years in terms of reducing the gap in between abilities of mankind and those of computers. Researchers are working on a variety of areas of the field in order to bring about amazing results. In this regard, the field of Computer Vision is only one among several. Among the objectives of this field is to enable machines to view and comprehend their environment in the same way that humans do, and to use this expertise for a variety of tasks such as image and video recognition, image processing & categorization , recommendation systems, etc. The advances in Computer Vision in regard with Neural Networks have been built and refined over time, mainly via the use of a single algorithm — the Convolutional Neural Network as the foundation.

When given an images as input, a Convolutional Neural Network (ConvNet) may assign significance to various features or objects in the image, and then distinguish between them. When contrasted to all other classification methods, the amount of pre-processing needed by a ConvNet is much less. While basic techniques need filters to be hand-engineered, ConvNets have the capability of learning these characteristics with sufficient training. In many ways, the design of a ConvNet is similar to the connection patterns of Neurons in the Structure Of the brain, and it was influenced by the structure of the Visual Cortex. Every single neuron in the human visual system responds to signals primarily in a certain area of the field of vision termed as Receptive Field. A group of similar fields may be used to fill the whole visual region by overlapping them.

## 3.4.1. Architecture of CNN

Every layer of a basic ConvNet converts one region of activation functions to another via the use of a differentiable function, and each layer of a simple ConvNet is composed of layers. Convolutional Network designs are built using three kinds of layers: the Convolutional Layer, the Pooling Layer, and the Fully-Connected Layer. These layers will be stacked together to create a complete ConvNet architecture.

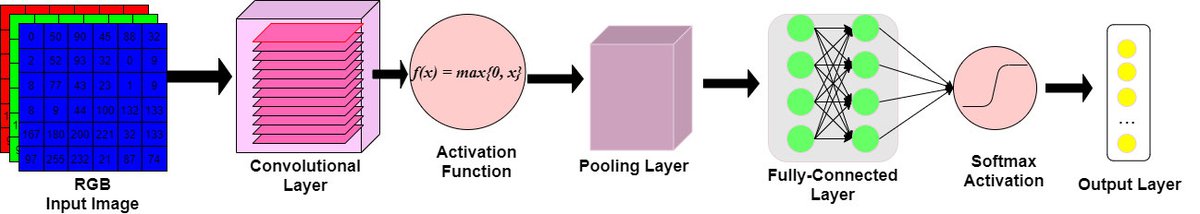


Figure 5 - Architecture of CNN

As mentioned above there are various layers in the architecture of a CNN. The layers can be described as follows:

1. Input Layer: The raw image pixels of the picture will be passed through the input layer.
2. Convolutional Layer: Convolution layer is the most essential building blocks of convolutional neural networks, and they are utilized in many different applications. Convolutional layers operate on the input by applying a convolution operation and transferring the output towards the next layer. A convolution is a mathematical operation that transforms all of the pixels through its receiving area it in to a numeric measure. For instance, if we perform a convolution to a picture, we will be reducing the size of the image while also condensing all of the data in the field into single pixel in the process. A vector is produced as the final outcome of the convolutional layer. Different types of convolutions may be used depending on the kind of issue we are attempting to solve and the type of the attributes we are attempting to learn from the data.
3. Pooling Layer: Pooling layers are a technique for down sample feature maps that summarizes the existence of features in regions of the feature map, allowing for more efficient sampling. Average pooling & max pooling are both often used pooling techniques that summaries the mean existence of a feature as well as the most active presence of a feature, respectively, in a dataset.
4. Fully connected Layer: Feed forward networks are what the Fully Connected Layer is all about. The Fully Connected Levels are the last few layers of the network hierarchy. A flattened version of the outputs from the last Pooling Layer is sent into the fully connected layer, which is then fed back into the fully connected layer.

The implementation of CNN is :



Test loss : 0.3171

Test Accuracy : 0.8348

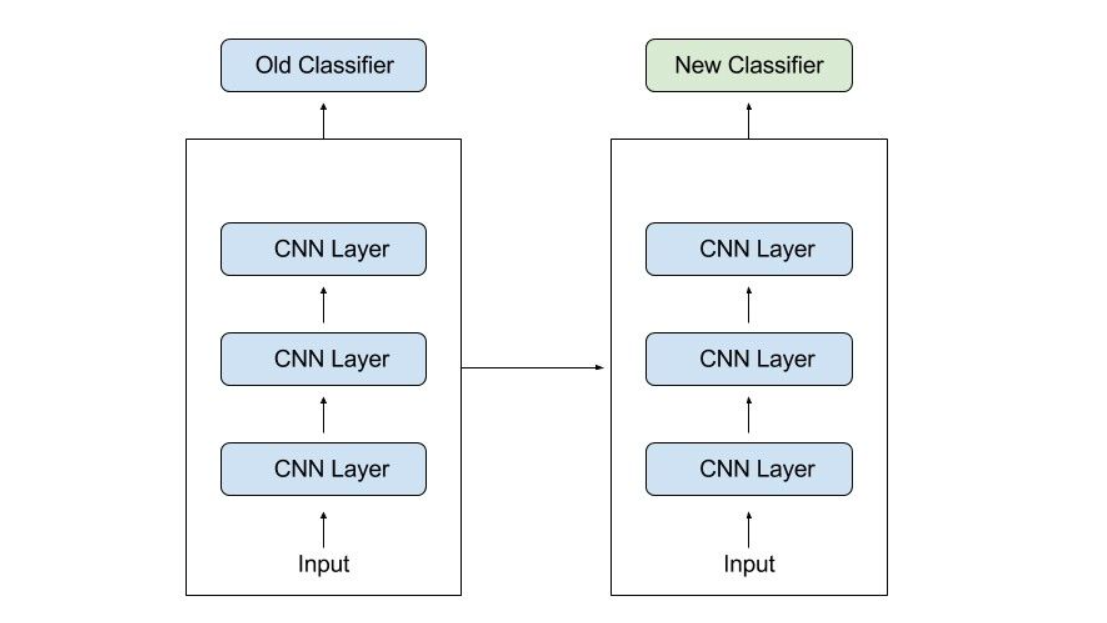
## 3.5 Transfer Learning

Transforming a model is trained on a huge database to another smaller one is at the heart of transfer learning. To use a CNN for object recognition, we freeze the network's early convolutional layers and train only the final few levels that make predictions. A pre-trained model is reused on a new problem through transfer learning (TL). Due to its ability to train convolutional neural networks with relatively little input, it's now quite popular among deep learning researchers. Since most real-world situations don't have millions of labelled data sets to train such complicated models, this is extremely helpful inside the data science field. Our goal is to provide you a better understanding of transfer learning. Transfer learning will be discussed, as well as some resources on pre-trained models will be provided.

## 3.5.1 Working of Transfer Learning

Brain networks are often used in computer vision to recognize edges, forms and task-specific properties in the early levels of the computer's vision system. The early to mid layers are utilized in transfer learning, and only the latter layers are retrained. It makes use of the labelled data from the task this was originally trained on.

Consider the case of a model that was trained to recognize a backpack in a picture and be used to detect sunglasses. For this reason, only the last layers will be retrained to learn how sunglasses differ from other objects. The model has already learned to distinguish objects in the prior levels.



As much information as feasible from the previous work is transferred to the current task when using transfer learning. Information can take a variety of shapes based on the situation and the available data. For example, the way models are put together may help us detect new items more quickly.

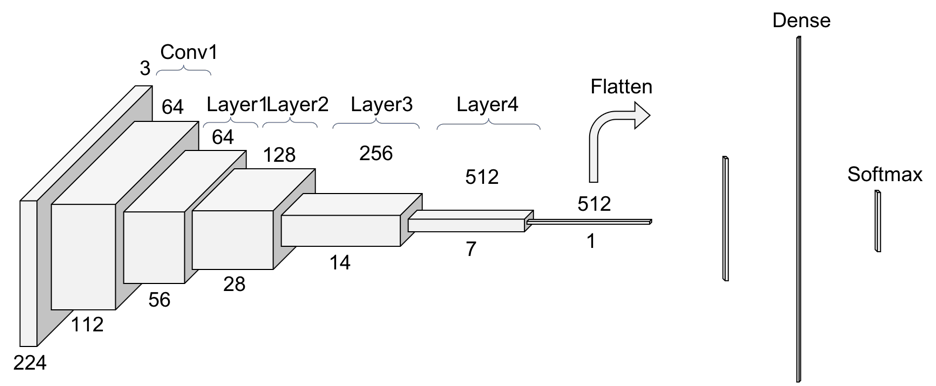
In this project I have implemented two models of transfer learning.

## 1. Resnet v2

It differs from regular sequential networks in that ResNet uses microarchitecture modules known as a network inside the architecture rather than sequential network such as OverFeat and VGG. New networks are often built from the ground up using microarchitecture, which is a collection of small building blocks. Based on Kaiming He's seminal work that was first published in 2015, this research shows how deep network can be train using regular SGD by modifying the residue module to employ identity mapping as proven in the previous work.

In spite of ResNet's greater depth, the real model weights are smaller because average global pooling instead of the convolution layer is employed, resulting in a smaller model size for ResNet50, which is more commonly used for 100MB's models.

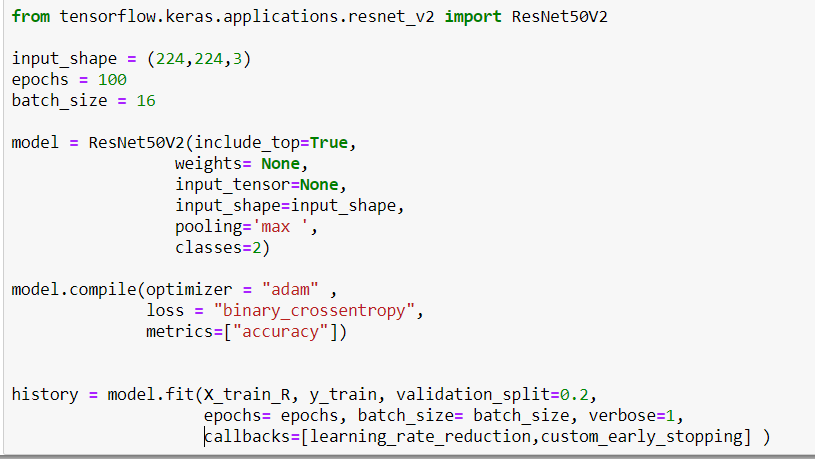
## Architecture :



These are the some important points of Resnet v2.

* It is based on the 'Residual Learning' idea. Residual Learning that is based on stacking residual blocks to improve performance of neural networks.
* Resnet stands for Residual Network, which is a network that facilitates Residual Learning. The number 50 denotes the number of layers. Resnet50 refers to a 50-layer residual network.
* Due to their inability to extract significant information from images, simpler Neural networks do not work well on dataset.
* As a result, the number of hidden layers will automatically grow in order to improve performance indicators like as accuracy and AUROC. However, when we add additional layers to our neural networks, the accuracy begins to saturate and ultimately decline. Residual Training aims to address this issue.
* Instead of attempting to learn some attributes, residual learning focuses on learning some residual. Residual could be simply defined as the removal of a feature learned from a layer's input. Shortcut connection (directly connecting the input of the nth layer to the input of some (n+x)th layer) are used by ResNet to accomplish this. It has been demonstrated that training this type of network is simpler than training simple cnn model, and that the issue of accuracy degradation has been handled.
* Inputs can propagate quicker through back propagation across layers with residual blocks, and Batch Normalization layers are used to speed it up runtime and prevent overfitting.
* The Vanishing Gradients issue is also addressed by ResNets. It prevents the gradients from rapidly decreasing to zero. A "Quick" or a Skip Connections in ResNet allows a gradients to be directly back-propagated to previous layers.

The Implementation of Resnet V2 for this project is :



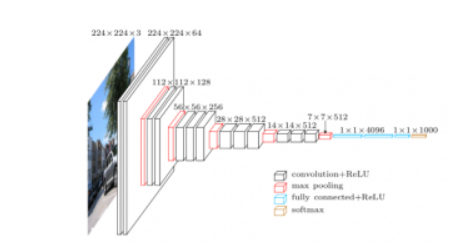
The accuracy on testing data is 72.72% and Testing loss is 1.55.

## 2. VGGNet

Karen Simonyan and Andrew Zisserman of the University of Oxford proposed the VGGNet architecture for a Convolutional Neural Network in 2014. The main focus of this research is on the accuracy of the cnn model as a function of its depth. Very Convolutional Neural Network for Wide Scale Image Processing is the title of the original study on VGGNet.

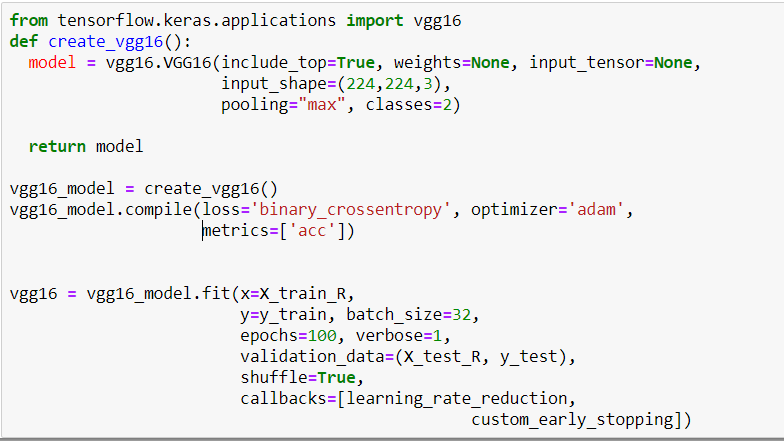
## VGGNet Architecture

When using VGG-based convNet, you'll need a 224\*224 Rgb as your input. The preprocessing layer subtracts the average image values determined for the complete ImageNet training set from the Image with pixels values ranging from 0–255.



These weight layers are applied to the preprocessed input photos. A series of convolutional layers is applied to the practice images. In the VGG16 architecture, there are a total of Thirteen convolution layer and three fully connected layers. Instead of using huge filters, VGG uses smaller ones (3\*3) that have more depth in their image processing capabilities. Using this method, the receptive field is identical to that of using only one 7 x 7 convolution layers.

The Implementation of VGG16 for this project is :

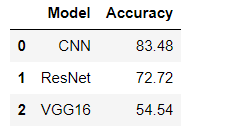


The Testing loss for VGG16 Network is 68.90 % and the accuracy on testing data is 54.54%.

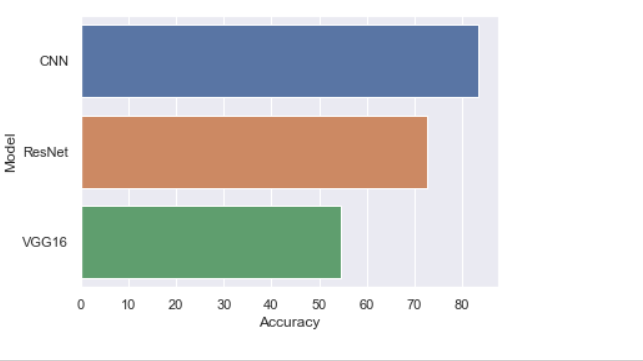
## 4. Evaluation Metrics

The evaluation metrics used to evaluate the CNN models and Transfer learning models is Accuracy. On the basis of training data, accuracy is the metric used to evaluate which model is the most effective at detecting patterns and connections among variables in a dataset.

After training the models on the training set using the accuracy as metric, I have further evaluated the models on the test set that was provided along with the dataset.



The bar graph of the accuracy is :



## 

## Conclusion

In this project we are trying to predict the Cancer using the image evidences that were provided in the dataset. Using this dataset, we have performed several pre-processing operations and trained CNN and Transfer learning with different architectures. After applying different architectures we can see that CNN gives higher accuracy with testing data and the testing loss is very less.

Upon comparing the performances of all the models that were trained as part of this project it has been observed that the performance of the seven layer CNN with 3x3 kernel size of convnets is more efficient on all the training and test sets. Various neural network strategies for skin cancer detection and classification were discussed in this systematic review research. These methods are all non-invasive. Preprocessing and picture segmentation are two phases in the detection of skin cancer. Following that, feature extraction and categorization are performed. The focus of this review was on CNNs and transfer learning for lesion picture classification. Each algorithm has its own set of benefits and drawbacks. The most important factor in achieving the best results is choosing the right classification technique. When it comes to classifying picture data, however, CNN outperforms other types of neural networks because it is more directly tied to computer vision than other types of neural networks.

Based on the analysis my findings from answering the research questions:

The fundamental advantage of CNN over its predecessors is that it discovers essential traits without the need for human intervention. Given a large number of photos of cats and dogs, it can learn the key characteristics of each class on its own. With the many CNN-based deep neural networks constructed, a considerable result was reached on the ImageNet Challenger, which is the most significant picture classification and segmentation challenge in the image analyzing area . In medical categorization, the CNN-based deep neural system is commonly utilized. It has been observed that CNN gives better accuracy than Transfer learning models.