© May 2022 | IJIRT | Volume 8 Issue 12 | ISSN: 2349-6002

Analysis and Implementation of Skin and Breast Cancer Prediction Using Deep Learning

Muskan Razdan[1], Nikhil Saini[1], Naman Kaushik[1], Mahesha A. M.[2]
[1]*CSE DEPARTMENT. JSS ACADEMY OF TECHNICAL EDUCATION, NOIDA*
[2]*Assistant Professor, Dept of CSE., JSS ACADEMY OF TECHNICAL EDUCATION, NOIDA*

Abstract- **In recent years, human infections have expanded unexpectedly everywhere. And one of those is Breast Cancer, Breast Cancer has expanded at a disturbing rate in the previous decade and this pattern of increment would keep on developing. Presently, there is a requirement for proficient message investigation and element extraction instruments to help with grouping, sharing, and recovering data on human illnesses overall and Breast Cancer specifically. Skin Cancer, the most well-known human danger, is essentially analyzed outwardly, starting with an underlying clinical screening and followed possibly by dermoscopic investigation, a biopsy and histopathological assessment. Mechanized arrangement of skin sores utilizing pictures is a difficult undertaking attributable to the fine-grained changeability in the presence of skin sores. Cancer Diagnosis Assistance Tool is designed for the diagnosis of certain types of cancer which will be an outcome of deep learning, data analytics and image processing. This Project will work as an assistance tool for the doctors and it will increase the accuracy of the diagnosis. Profound learning with Convolutional Neural Networks has arisen as one of the most impressive AI apparatuses in Image order, outperforming the exactness of practically any remaining conventional arrangement techniques and surprisingly human capacity. The convolutional cycle can work on a picture containing a huge number of pixels to a bunch of little component maps, subsequently lessening the element of info information while holding the main differential highlights. Therefore in our research, CNN is used to classify the images. Our research is based on the images and CNN is the most popular technique to classify the images. The proposed system is found to be successful, achieving results with 87% accuracy, which could reduce human mistakes in the diagnosis process. Moreover, our proposed system achieves accuracy higher than the 78% accuracy of machine learning (ML) algorithms. The proposed framework, in this way, further develops accuracy by 9% above outcomes from AI (ML) calculations.**

Index Terms- **Breast Cancer, Skin Cancer, Deep Learning, Convolutional Neural Network, ANN.**

I. INTRODUCTION

The current is an era of technologies, and in this era everything is getting automated and digitalized and so are our healthcare systems. Many new advancements have taken place in the medical world in the past few years. This Breast Cancer And Skin Cancer Prediction Using Deep Learning will be our contribution to medicare that can help Doctors to identify the risk of certain diseases. Bosom Cancer is one of the main diseases created in numerous nations including India. If a woman gets diagnosed in an early stage the preservenace rate can increase upto 97%. Reportedly, in the last few decades the number of deaths due to breast cancer has increased rapidly. According to WHO, 627,000 women died from breast cancer in 2018. Breast cancer is the main problem that spreads everywhere in the world but is mostly found in the United States of America. There are four types of breast cancer. The first type of cancer is Ductal Carcinoma in Situ that is found in the coating of breast milk ducts and is pre-stage breast cancer. The second type of breast cancer is the most popular disease and contains up to 70-80% of diagnoses. The third type of breast cancer is Inflammatory breast cancer which is forcefully and quickly developing breast cancer in this disease cells penetrate the skin and lymph vessels of the breast. The fourth type of breast cancer is Metastatic breast cancer which spreads to other parts of the body.

Skin cancer is the most common malignancy in Western countries, and melanoma specifically accounts for the majority of skin cancer-related deaths worldwide [11]. In recent years, many skin cancer classification systems using deep learning have been developed for classifying images of skin tumors, including malignant melanoma (MM) and other skin cancer [10]. There are reports that their accuracy was at the same level as or higher than that of dermatologists. There are three main types of skin cells: Squamous, Basal, and Melanocytes. Skin cancer is commonly caused due to skin exposure to sunlight(UV Rays). Older adults and people with suppressed immune systems have a high risk of dying from skin cancer. The key challenge with its treatment is early detection.

As a result, in addition to pharmaceutical therapies, some Data Science solutions must be included to address these death-causing concerns. The goal of this study is to identify which traits are most useful in predicting whether a cancer is malignant or benign, as well as to look for general trends that might help us choose models and hyperparameters.

The goal is to find out whether the patient is suffering from breast cancer or not. To achieve this we have used deep learning methods to fit a function that can predict the discrete class of new input.

II. CANCER DIAGNOSIS AND ANALYSIS

Cancer diagnosis involves several steps, some of which might result in incorrect test findings. In certain circumstances, cancer may be preventable.

- **Physical examination:** The doctor searches for anomalies that might suggest the presence of cancer, such as changes in skin color or organ enlargement.
- **Laboratory testing**, such as urine and blood tests, can aid in the detection of anomalies that may be caused by cancer.
- **Imaging Tests** include a computed tomography (CT) scan, bone scan, magnetic resonance imaging (MRI), positron emission tomography (PET) scan, ultrasound, and X-ray, among others, and allow the doctor to inspect your bones and internal organs. ultrasound and X-ray, among others.
- **Biopsy.** During a biopsy, your doctor takes a sample of cells for laboratory testing. Doctors examine cell samples under a microscope in the laboratory. Normal cells have identical diameters and are organized logically. Cancer cells have a disorganized appearance, with various sizes and no obvious structure. [4]

  Because various factors, such as differences in your body or even what you consume, might impact test outcomes, test results must be carefully understood. It's also worth remembering that noncancerous illnesses can occasionally lead to aberrant test findings. And, in other circumstances, cancer may be present despite normal blood test findings.

III. MATERIAL AND METHODS

We looked through Google Scholar, Elsevier, PubMed, Research gate and took a portion of the important research deals with skin characterization utilizing deep learning concepts , using Convolutional Nueral Networks, furthermore, we did an audit on the exhibition measurements, furthermore, discoveries on how involving profound learning in skin malignant growth characterization has further developed exactness on characterizing skin malignant growth contrasted with a dermatologist.

We recognize that the possibility of survival increases drastically for the majority of patient groups if cancer is detected at an early stage. Existing screening programs like breast, cervical and colorectal cancer have saved a great many lives, but they lack sensitivity.[5] To increase the biological understanding and awareness of .ealy cancer and enhance the ability of our heallthcare system to diagnose this vicious disease research play a key role and with that we can look up the technologies that might help in the finding an optimized solution. High-quality research focused on the early detection of cancer is currently being carried out in several locations. Cancer cells come in a variety of sizes and shapes. It might be difficult to identify early on.[3]

Dataset

The Histopathological dataset from Kaggle was used for the proposed system. This data set consists of both benign and malignant images. The careful observation was ensured during splitting; the dataset was divided into validation data and testing data belonging to the same distribution to well represent the model's generalized results. For learning indicators like weights and biases, training data is important, while validating data is essential for model verification and how exactly the model simplifies, thus tuning hyperparameters like learning rate and decay to boost the result of the model.[9] A model's final output comes from precise work on the test results. To hold each pixel in the same range and prevent bias, normalization has to be done on the whole image.

Performance Metrics

Model Evaluation is performed by the following metrics [12].

Specificity: It gives what fraction of all negative classes are correctly identified as negative my model classifier. Also called True Negative Rate.

SP = TN/ (TN + FP) Sensitivity: It is the ability of the process to correctly identify the disease condition or situation.

SE = TP/ (TP + FN) ROC AUC: Area under receiver operating characteristic. It is the probability that the classifier will identify TPR against FPR. It is the graph between true positive rate vs false positive rate.

Precision: It gives what fraction of all classes that are correctly predicted positive, are actually positive. **PREC = TP / (TP + FP) Negative Predictive Value (NPV):** It is the probability of the inputs that tested negative, truly does not have the disease.

NPV = TN / (TN + FN) Positive Predictive Value (PPV): It is the probability that the classes that are tested positive, truly have the disease.

PPV = TP/ (TP + FP) Dice coefficient: It gives the overlap measure between the automatic and the ground truth segmentation. It is also called an overlap index.

DC = 2TP/ (2.TP + FN + FP) Where True negative (TN): negative class is correctly predicted by the classifier model

True positive (TP): positive class is correctly predicted by the classifier model

False negative (FN): negative class is incorrectly predicted by the classifier model False positive (FP): positive class is incorrectly predicted by the classifier model.

Preprocessing of Data

In the pre-processing stage, the median filter will be used for Adaptive contrast enhancement(AHE). The AHE is capable of improving local contrast and bringing out more details in the image. Therefore, contrast limited adaptive histogram equalization which is a type of AHE will be used to improve the contrast in mammogram images of breast cancer images.[2]

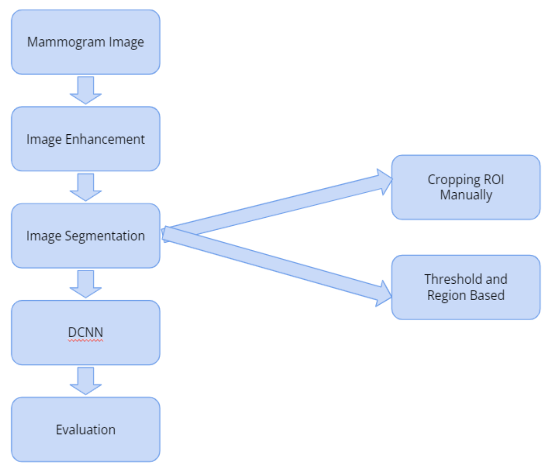


Fig. 1. Block diagram of the proposed breast cancer model

Figure 1 depicts the proposed Deep Convolutional Neural Network technique for detecting breast cancer based on mammography images. Image segmentation will be used after picture enhancement to divide an image into portions with comparable traits and qualities. The first technique will use circular contours to launch the Region of Interest pooling. The threshold and region-based approaches will be utilized to resolve the Region of interest pooling in the second method. The Region of interest pooling will be classified using Deep CNN.

A. CNN Architecture:

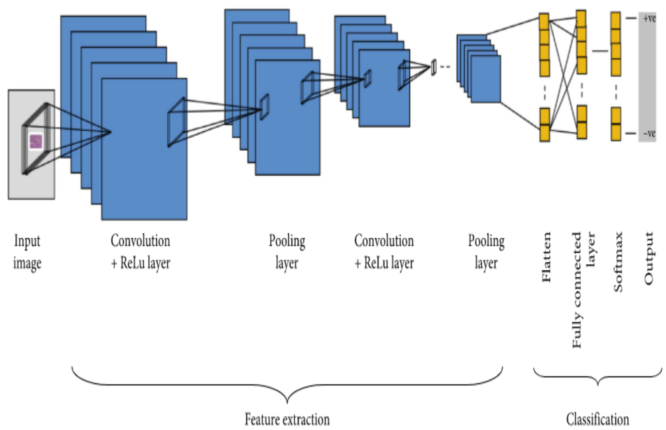


Fig 2: CNN Architecture

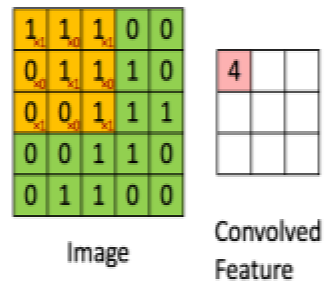
Convolutional neural networks are the leading architecture in deep learning that is used to solve an image classification problem.[1] The process of building a convolutional neural network always involves 4 major steps:

- Convolution
- Pooling
- Flattening
- Fully connected layer

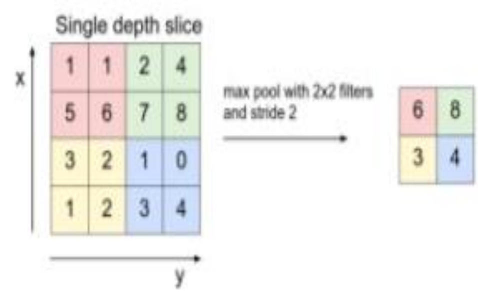
Types of layers:

The name of a layer comes from the fact that all neurons in that layer perform comparable mathematical processes.

1. **Convolution layer:** Convolution is the mathematical operation that is used in image processing to filter signals, find patterns in signals, etc. All neurons in this layer perform convolution on inputs. The most important parameter in a convolutional neuron is the filter size. We shall slide convolution filters over the whole input image to calculate this output across the image and here we slide our window by 1 pixel at a time this number is called Stride. Typically we use more than 1 filter in one convolution layer.

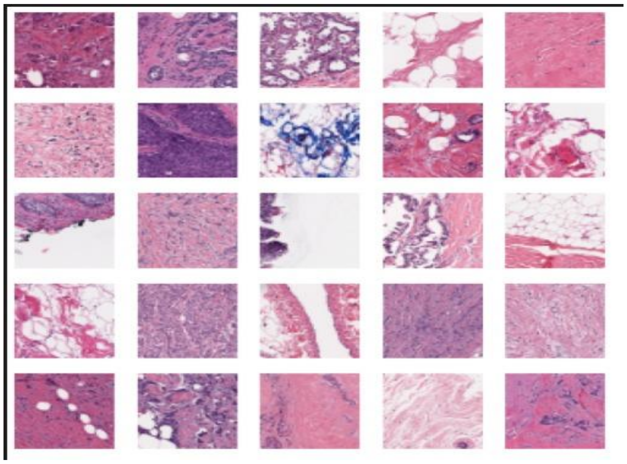


2. **Pooling layer:** To lower the spatial size, a pooling layer is usually utilized right after the convolutional layer (only width and height, not depth). As a result, the number of parameters is reduced, and the computing time is lowered. Furthermore, a smaller number of parameters prevents overfitting. The most popular type of pooling is Max pooling, which involves taking a 3X3 filter and doing the maximum operation on the 3X3 sized portion of the picture.



3. **Fully Connected Layer:** If each neuron in a layer receives input from all the neurons in the previous layer, then this layer is called a fully connected layer. The output of this layer is computed by matrix multiplication followed by bias offset.[6]

Training and Testing the CNN using original data



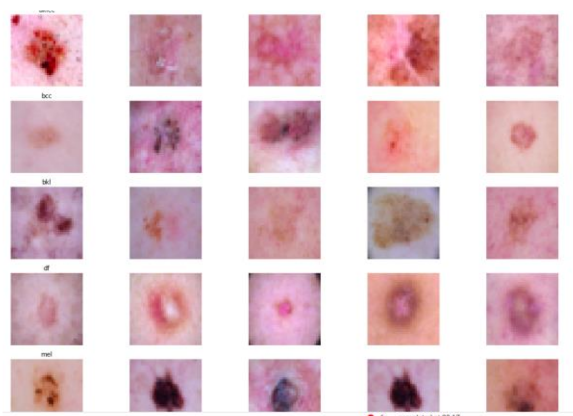
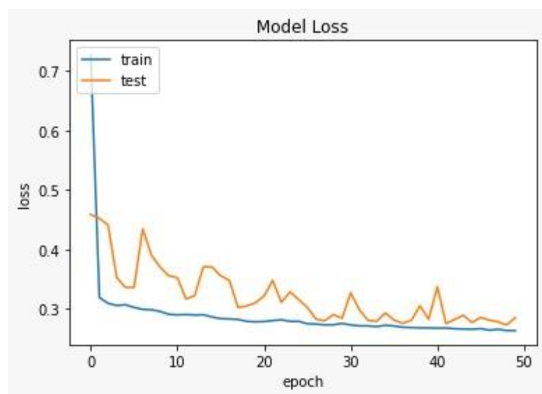


Fig.3.Sample images from Histopathological dataset.
i) Breast Cancer
ii) Skin Cancer

IV. RESULT

Satisfactory results have been obtained using the CNN-based proposed breast cancer detection method. From the studies that are reviewed, CNN was the best in performing classification of all other architectures.[4] Using archives that mostly contain skin lesions from light-skinned persons is a significant problem in our study. For example, the images of ISIC are mainly from the United States, Australia, and Europe. CNN must be trained to abstract from diverse skin tones to produce reliable classification results for dark-skinned persons. This may be accomplished by taking into account dark-skinned photos. Clinical data of various ages, picture size, gender, and skin type can be used as inputs for classifiers to improve classification quality. Training and testing were done through two methods, in the first method the dataset was divided into two classes named normal and abnormal. While the second method included a further subdivision of the abnormal classes, which included six types of abnormalities found in breasts such as asymmetry, calcification, spiculated masses, circumscribed masses, architectural distortion, and miscellaneous. Miscellaneous photos were ones in which it was unclear if the images were benign or malignant. The purpose of pre-processing is to improve the performance and learning speed of neural networks. The accuracy of raw pictures acquired with various filter sizes in CNNs. When the model parameters are learned and fixed with no further learning, accuracy is calculated.



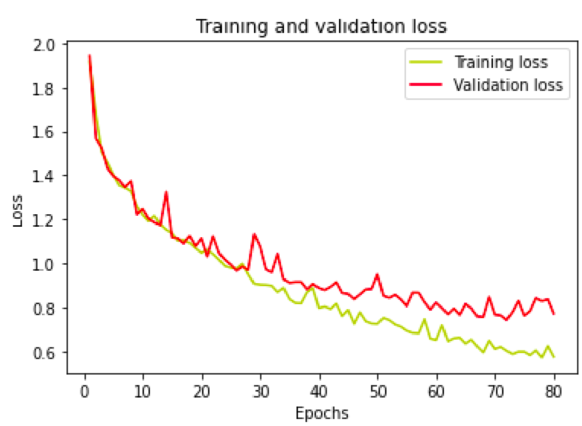


Fig 4. Loss Graph: The loss is minimized in the training Process
i) Breast Cancer
ii) Skin Cancer

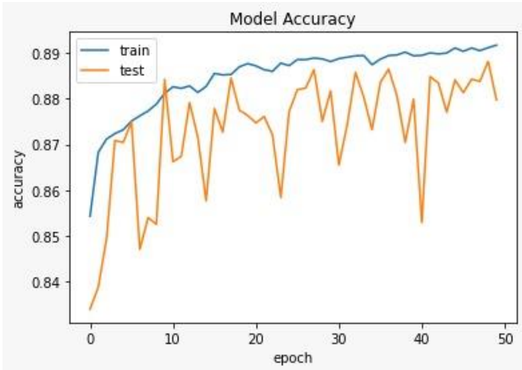
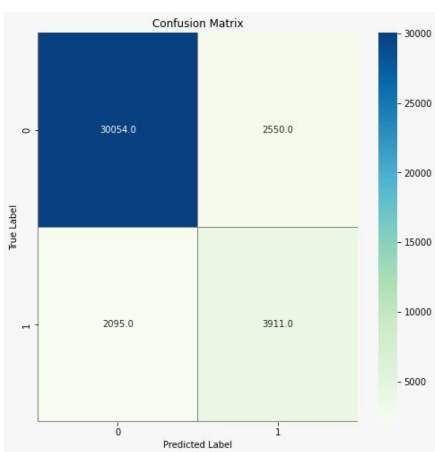


Fig 5. Accuracy Graph: The Accuracy of the system gets maximized with training

i) Breast Cancer
ii) Skin Cancer



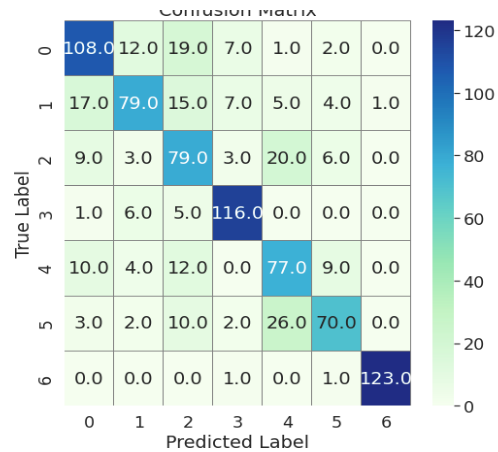
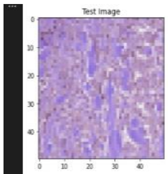


Fig. 6. Confusion Matrix: Describes the Performance of Model

V. CONCLUSION

Breast cancer detection is a challenging problem because it is the most popular and harmful disease. Breast cancer is growing every year and there is less chance to recover from this disease. For detection of breast cancer, machine learning and deep learning techniques are used. It is concluded that machine learning gives better results on linear data only therefore, for the classification of the breast cancer images data, a deep learning based technique CNN is used. CNN mostly works on the images dataset.

In this work, we presented a deep learning approach for the detection of breast cancer using mammograms. The proposed approach is developed following the development of convolution neural networks and it demonstrates how robust deep learning is in this application. There are several potential variations of the proposed network architecture that can be investigated and validated as a future work. The proposed method can lead to better performance of clinical use of breast cancer detection especially in early stages.
The proposed system is found to be 87.97% Accurate with a loss of 0.2852.




```
input1 = X_test[index1:index1+1]
print('Input Index =',index1)
```

```
Input Index = 5
```

```
cnn_pred1 = forest.predict(input1)[0].argmax()
label1 = y_test[index1].argmax()
```

```
print('Predicted Value using  cnn model',cnn_pred1)
print("\nTrue Value",label1)
```

```
Predicted Value using  cnn model 1
True Value 1
```

Fig 7. Result of a Sample Image.
Here True Value 1 Means that the Patient is Suffering From Breast Cancer

	precision	recall	f1-score	support
0	0.93	0.92	0.93	32604
1	0.61	0.65	0.63	6006
accuracy			0.88	38610
macro avg	0.77	0.79	0.78	38610
weighted avg	0.88	0.88	0.88	38610

Fig 8 Classification Report of the Model
Precision: Quantifies the number of Positive class predictions
that actually belong to Positive Class
Recall: The Model's Ability to detect Positive Samples
F1-Score: The Harmonic Mean of Precision and Recall

```
[ ] loss,accuracy=model.evaluate(X_test,y_test)

1207/1207 [==============================] - 6s 5ms/step - loss: 0.2852 - accuracy: 0.8797
```

Fig 9. Loss and Accuracy evaluation of the Model

REFERENCES

[1] Dina A. Ragab, Maha Sharkas, Stephen Marshall and Jinchang Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines", Published January 28, 2019, View at:https://peerj.com/articles/6201/

[2] M. Masud, A. E. Eldin Rashed, and M. S. Hossain, "Convolutional neural network-based models for diagnosis of breast cancer," Neural Computing and Applications, vol. 5, 2020. View at: Publisher Site | Google Scholar

[3] G. Muhammad, M. S. Hossain, and N. Kumar, "EEG-based pathology detection for home health monitoring," IEEE Journal on Selected Areas in Communications, vol. 39, no. 2, pp. 603–610, 2021. View at: Publisher Site | Google Scholar

[4] Saad Awadh Alanazi, M. M. Kamruzzaman, Md Nazirul Islam Sarker, Madallah Alruwaili, Yousef Alhwaiti, Nasser Alshammari, Muhammad Hameed Siddiqi, "Boosting Breast Cancer Detection Using Convolutional Neural Network", Journal of Healthcare Engineering, vol. 2021, Article ID 5528622, 11 pages, 2021. https://doi.org/10.1155/2021/5528622

[5] Rakha, E. A. et al. "Visual histological assessment of morphological features reflects the underlying molecular profile in invasive breast cancer: a morpho-molecular study. Histopathology". 77, 631–645 (2020).

[6] Shwetha K, Spoorthi M, Sindhu S S,Chaithra D , "Breast Cancer Detection Using Deep Learning Technique". International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181

[7] Syed Jamal Safdar Gardezi, Ahmed Elazab, Baiying Lei, Tianfu Wang, "Breast Cancer Detection and Diagnosis Using Mammographic Data", J Med Internet Res. 2019 Jul; 21(7): e14464.

[8] Li Shen , Laurie R. Margolies, Joseph H. Rothstein, Eugene Fluder, Russell McBride & Weiva Sieh,"Deep Learning to Improve Breast Cancer

Detection on Screening Mammography", Nature Research 2019.

[9] Deepak Motwani, Kapil Sharma , Ashok Shrivastava, "Feasible Study of Breast and Lungs Cancer Detection using Machine learning ", Engineering and Technology Journal for Research and Innovation (ETJRI) ISSN 2581-8678, Volume III, Issue I, Dec 2020.

[10] Shunichi Jinnai , Naoya Yamazaki , Yuichiro Hirano , Yohei Sugawara , Yuichiro Ohe and Ryuji Hamamoto, "The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning", mdpi, 29 July 2020.

[11] Indian Cancer Society. 2020. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancerfacts-and-figures/2020/cancer-facts-and-figures-2020.pdf (accessed on 1 June 2020)

[12] Ravi Manne,| Snigdha Kantheti, Sneha Kantheti, "Classification of Skin cancer using deep learning, ConvolutionalNeural Networks - Opportunities and vulnerabilities- A systematic Review" , International Journal for Modern Trends in Science and Technology, 18 November 2020