A Comparison of Different CNN Frameworks for Skin Cancer Detection

A Project Report

Submitted by,

Ch. Aaron John - 18BCE0204
R Ravi Maithrey - 18BCE0459
Ippili Akarsh - 18BCE2523

Under the guidance of,

Dr. Gopinath M. P.
SCOPE, VIT, Vellore

CSE4020 - Machine Learning

Abstract

Skin cancer is a very dangerous disease to contract. Yet many people are unsure if the growths on their bodies are natural or carcinogenic. This is because figuring out whether a growth is carcinogenic or not is a matter of delicate analysis To the human eye, unless highly trained, it is very hard. But a computer system can be trained to do this same task and reduce both the work to be done and provide better results. A neural network when used in this scenario will provide most flexibility because they can be taught without regarding parameters. This allows us to make significant advances in identifying skin cancers before they pose major threats. The main focus of this paper would be to identify the different frameworks within neural networks which can be leveraged for this purpose and then compare them to identify the best.

Keywords: Skin cancer detection, Neural Networks, Convolutional Neural Networks, CNN Frameworks

Introduction

Skin cancer is a common form of cancer, and early detection increases the survival rate. Skin cancer is an alarming disease for mankind. The necessity of early diagnosis of the skin cancer have been increased because of the rapid growth rate of Melanoma skin cancer, it's acute; high treatment costs, and death rate. This cancer cells are detected manually and it takes time to cure in most of the cases.

According to the world cancer report, the primitive reason of melanoma is ultra violet light exposure in those people who have low level of skin pigment. The UV ray can be from the sun or any other sources and approximately 25% of these malignant skin cells can form moles, which are easy to observe but cannot be perfectly stated whether they are benign or malignant without advanced specification techniques from the diagnostic side.

To solve this problem, we turn to neural networks. Neural Network algorithm is utilized to detect the benign and malignant. This framework is based on learning the images that are captured with imaging device to and out whether it is benign or malignant. Not all neural networks work well with this type of scenario though. Here, we need to be able identify features of images, and these images need to be read and understood with all the contextual information.

For this reason we use a special type of neural networks called Convolutional Neural Networks They allow us to retain all the contextual information in the data which can be sent in as images In addition to that, they preserve the dependencies in the image and carry it on to the calculation as well. Convolutional Neural Network (CNN) is a type of neural network which is used in signal and image processing.

Literature Survey

Object Detection Using Convolutional Neural Networks

In this paper, Convolutional Neural Networks (CNN) is used to detect objects in the environment. Two state of the art models are compared for object detection, Single Shot Multi-Box Detector (SSD) with MobileNetV1 and a Faster Region-based Convolutional Neural Network (Faster-RCNN) with InceptionV2. Result shows that one model is ideal for real-time application because of speed and the other can be used for more accurate object detection.

Applications of Machine Learning in Cancer Prediction and Prognosis

A broad survey of the different types of machine learning methods being used, the types of data being integrated and the performance of these methods in cancer prediction and prognosis. A number of trends are noted,including a growing dependence on protein biomarkers and micro-array data, a strong bias towards applications in

prostate and breast cancer, and a heavy reliance on "older" technologies such artificial neural networks (ANNs) instead of more recently developed or more easily interpretable machine learning methods.
A number of published studies also appear to lack an appropriate level of validation or testing. Among the better designed and validated studies it is clear that machine learning methods can be used to substantially (15-25%) improve the accuracy of predicting cancer susceptibility, recurrence and mortality.

Boosting Breast Cancer Detection Using Convolutional Neural Network

A convolutional neural network (CNN) method is proposed in this study to boost the automatic identification of breast cancer by analyzing hostile ductal carcinoma tissue zones in whole-slide images (WSIs). The paper investigates the proposed system that uses various convolutional neural network (CNN) architectures to automatically detect breast cancer, comparing the results with those from machine learning (ML) algorithms.

Review of deep convolution neural network in image classification

This paper first introduces the rise and development of deep learning and convolution neural network, and summarizes the basic model structure, convolution feature extraction and pooling operation of convolution neural network. Then, the research status and development trend of convolution neural network model based on deep learning in image classification are reviewed, which is mainly introduced from the aspects of typical network structure construction, training method and performance. Finally, some problems in the current research are briefly summarized and discussed, and the new direction of future development is forecasted.

Predicting breast cancer recurrence using effective classification and feature selection technique

This paper aims at finding breast cancer recurrence probability using different data mining techniques. We also provide a noble approach in order to improve the accuracy of those models. Cancer patient's data were collected from Wisconsin dataset of UCI machine learning Repository. This dataset contained total 35 attributes in which we applied Naive Bayes, C4.5 Decision Tree and Support Vector Machine (SVM) classification algorithms and calculated their prediction accuracy. An efficient feature selection algorithm helped us to improve the accuracy of each model by reducing some lower ranked attributes. Not only the contributions of these attributes are very less, but their addition also misguides the classification algorithms. After a careful selection of upper ranked attributes the papers found a much improved accuracy rate for all three algorithms.

Skin cancer diagnosis based on optimized convolutional neural network

In this paper, a new image processing based method has been proposed for the early detection of skin cancer. The method utilizes an optimal Convolutional neural network (CNN) for this purpose. In this paper, improved whale optimization algorithm is utilized for optimizing the CNN. For evaluation of the proposed method, it is compared with some different methods on two different datasets. Simulation results show that the proposed method has superiority toward the other compared methods.

Computer Aided Detection of Skin Cancer

This paper sees how with the advancement of technology, early detection of skin cancer is possible. One such technology is the early detection of skin cancer using Artificial Neural Network. The diagnosing methodology uses Image processing techniques and Artificial Intelligence. The dermoscopy image of skin cancer is taken and it is subjected to various pre-processing for noise removal and image enhancement. Then the image is undergone image segmentation using Thresholding.

There are certain features unique for skin cancer regions. Such features are extracted using feature extraction technique -2D Wavelet Transform method. These features are given as the input nodes to the neural network. Back-Propagation Neural (BPN) Network is used for classification purpose. It classifies the given data set into cancerous or non-cancerous.

Methodology

Convolutional Neural Networks - CNNs

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

The role of the ConvNet is to reduce the images into a form which is easier to process, without losing features which are critical for getting a good prediction.

The objective of the Convolution Operation is to extract the high-level features such as edges, from the input image. ConvNets need not be limited to only one Convolutional Layer. Conventionally, the first ConvLayer is responsible for capturing the Low-Level features such as edges, color, gradient orientation, etc. With added layers, the architecture adapts to the High-Level features as well, giving us a network which has the wholesome understanding of images in the dataset, similar to how we would.

The basic CNN architecture includes Convolution layer, Pooling layer, ReLU layer and a fully connected layer.

Convolution Layer: the real power of deep learning, especially for image recognition, comes from convolutional layers. It is the first and the most important layer. In this layer, a CNN uses different filters to convolve the whole image as well as the intermediate feature maps, generating various feature maps. Feature map consists of a mapping from input layers to hidden layers.

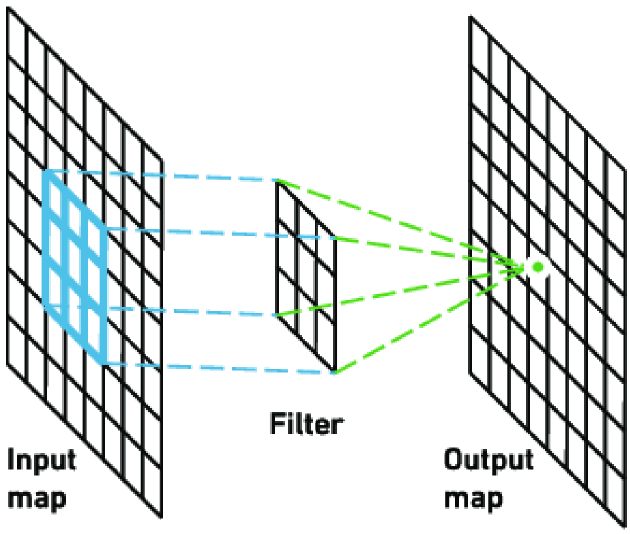


Figure 1: Typical Function of the Convolutional Layer

Pooling Layer: Similar to the Convolutional Layer, the Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data through dimensionality reduction. Furthermore, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model.

There are two types of Pooling: Max Pooling and Average Pooling. Max Pooling returns the maximum value from the portion of the image covered by the Kernel. On the other hand, Average Pooling returns the average of all the values from the portion of the image covered by the Kernel.

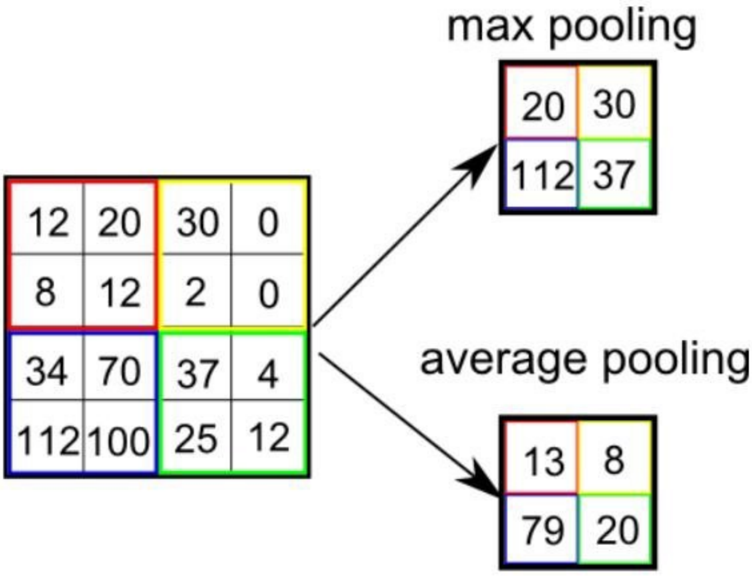


Figure 2: The types of pooling done in pooling layer

ReLU Layer: it is the Rectified Linear Units Layer. This is a layer of neurons that applies the non-saturating nonlinearity function or loss function:

$$f(x) = max(0, x)$$

It yields the nonlinear properties of the decision function and the overall network without affecting the receptive fields of the convolution layers. Pooling layer: its task consists to simplify or reduce the spatial dimensions of the information derived from the feature maps.

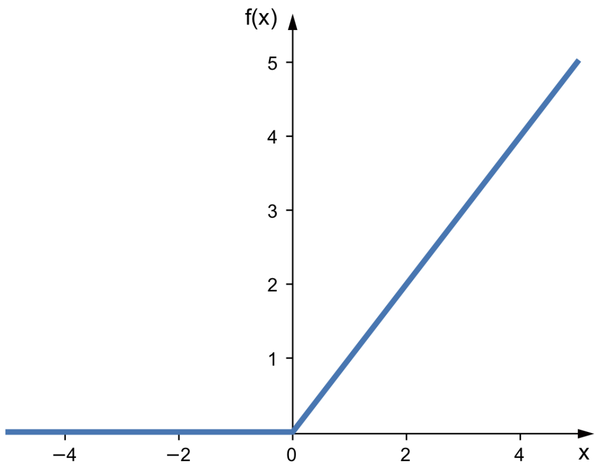


Figure 3: Typical Function of the Convolutional Layer

Fully Connected Layer:Adding a Fully-Connected layer is a (usually) cheap way of learning non-linear combinations of the high-level features as represented by the output of the convolutional layer. The Fully-Connected layer is learning a possibly non-linear function in that space. It helps us in gaining an output that is more efficient than without using this fully connected layer.

With the addition of this layer as well, the final CNN structure is as follows.

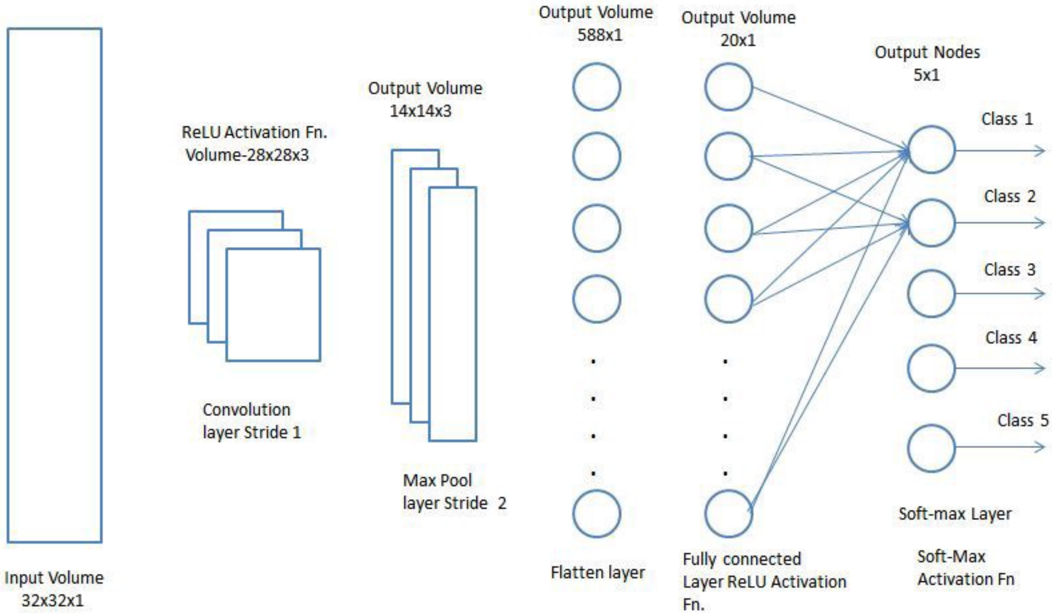


Figure 4: The full structure of the CNN

ResNet50

ResNet, short for Residual Networks is a classic neural network used as a backbone for many computer vision tasks. ResNet-50 is a convolutional neural network that is 50 layers deep. The network has an image input size of 224-by-224. The fundamental breakthrough with ResNet was it allowed us to train extremely deep neural networks with 150+layers successfully.

However, increasing network depth does not work by simply stacking layers together. Deep networks are hard to train because of the notorious vanishing gradient problem — as the gradient is back-propagated to earlier layers, repeated multiplication may make the gradient extremely small. As a result, as the network goes deeper, its performance gets saturated or even starts degrading rapidly.

ResNet first introduced the concept of skip connection. The diagram below illustrates skip connection.

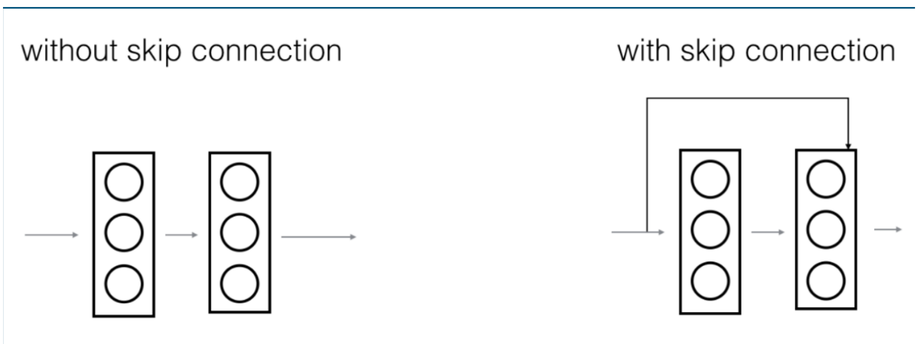


Figure 5: Skip connections in the ResNet

On the right we still stack convolution layers as before but we now also add the original input to the output of the convolution block. This is called skip connection.

There are two reasons why skip connections work here:

- They mitigate the problem of vanishing gradient by allowing this alternate shortcut path for gradient to flow through
- They allow the model to learn an identity function which ensures that the higher layer will perform at least as good as the lower layer, and not worse

Infact since ResNet skip connections are used in a lot more model architectures like the Fully Convolutional Network (FCN) and U-Net. They are used to flow information from earlier layers in the model to later layers. In these architectures they are used to pass information from the downsampling layers to the upsampling layers.

ResNet50 - Implementation

This architecture model for the CNN is implemented by way of sklearn which takes care of the image processing steps, along with opencv

Following that, a deep CNN is constructed with the help of keras module within tensorflow library

This keras allows us to have a deeper control over each layer of the neural network allowing us to fine-tune both the parameters and the hyper-parameters.

When defining the neural network itself the choice of the optimiser, activation function and the initial layer makes a huge difference in the way our neural works.

In our scenario, we use a regular initial layer, use RELU activation layer and finally use the ADAM gradient descent optimiser.

Efficientnet B0

EfficientNet-B0 is the baseline network developed by AutoML MNAS. EfficientNet is a convolutional neural network architecture and scaling method that uniformly scales all dimensions of depth/width/resolution using a compound coefficient. While scaling individual dimensions improves model performance, we observed that balancing all dimensions of the network—width, depth, and image resolution—against the available resources would best improve overall performance.

A comparison of the way different scaling methods effect the network are shown below

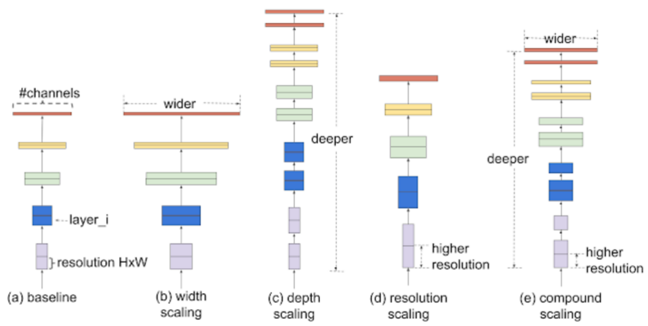


Figure 6: A comparison of the different scaling methods in use

EfficientNet scaling method uniformly scales network width, depth, and resolution with a set of fixed scaling coefficients. The compound scaling method is justified by the intuition that if the input image is bigger, then the network needs more layers to increase the receptive field and more channels to capture more fine-grained patterns on the bigger image.

EfficientNets, as the name suggests are very much efficient computationally and also achieved state of art result on ImageNet dataset which is 84.4% top-1 accuracy.

EfficientNet B0 - Implementation

This architecture model for the CNN is implemented by way of sklearn which takes care of the image processing steps, along with PILLOW. Following that, a deep CNN is constructed with the help of keras module within tensorflow library

This keras allows us to have a deeper control over each layer of the neural network allowing us to fine-tune both the parameters and the hyper-parameters. When defining the neural network itself the choice of the optimiser, activation function and the initial layer makes a huge difference in the way our neural works.

The first part of this process is to perform image segmentation, that is, to remove the unneeded parts of the images and to isolate those parts of the image which contain only the required classification parts

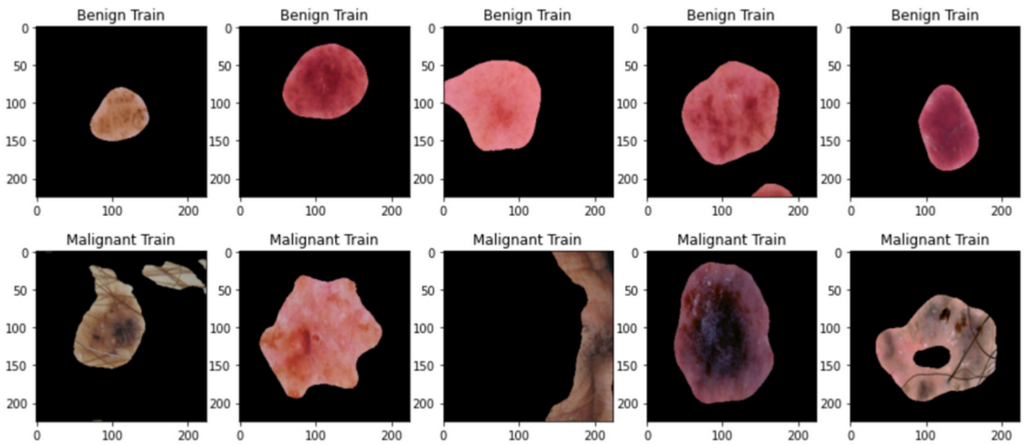


Figure 7: Segmentation of the images

Following this step we can do the pre-processing steps including the train, test and cross-validation split. We then rescale these images to make them fit to the input vector taken in by the EfficientNet architecture.

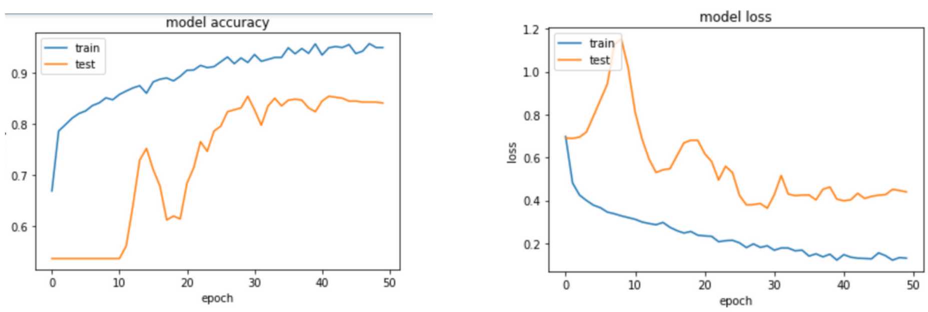
Then we can proceed with building our model. We define the layers of the neural network and then add their respective activation function. Since, the end result has to be us making sure we have only two outputs, either 1 or 0, so we use a softmax function on the last layer.

Results

ResNet50

After we finish training and then introduce our test set into the model, we can figure out how well our model did.

The final accuracy comes down to 85% with the model accuracy and loss shown below

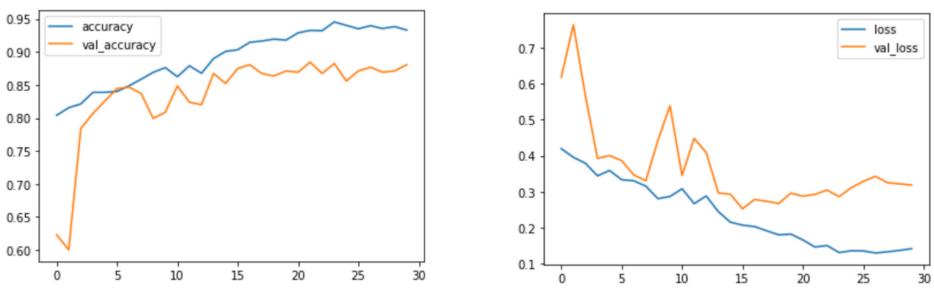


(a) Model Accuracy (b) Model Loss

Figure 8: The results of ResNet in classification

EfficientNet B0

While training our model we can see from the graphs below how loss and accuracy changes over the course of different batches



(a) Model Accuracy (b) Model Loss

Figure 9: Result of EfficientNet in classification

We can then proceed with using our trained model to give results. This gives us a confusion matrix with the following results

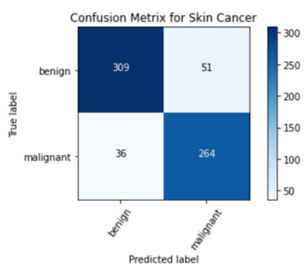


Figure 10: Confusion Matrix

Conlcusion

From the graphs and the accuracy scores shown above we can conclude that Efficientnet performs better than Resnet. This is because while both offer a similar value of accuracy, Efficientnet architecture only needs an average of 12 epochs to reach this accuracy value. Resnet on the other hand needs a minimum of 25-30 epochs for the same accuracy using the same parameters and the same dataset. Therefore, we can conclude that Efficientnet performs better while using lesser time to give the same performance.

For the purposes of classification of skin cancer, we require a model which is able give out the proper classification of malignant or benign with as little processing power and overheads as possible. For this reason, the classification using CNNs was performed and within the existing framework of architectures, a comparison was performed and then the most appropriate architecture for this model was determined.

References

[1] R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra, and J. M. Z. Maningo, "Object detection using convolutional neural networks," in *TENCON 2018 - 2018 IEEE Region 10 Conference*, pp. 2023–2027, 2018.

[2] J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and prognosis," *Cancer Informatics*, vol. 2, p. 117693510600200, 2006.

[3] S. A. Alanazi, M. M. Kamruzzaman, M. N. I. Sarker, M. Alruwaili, Y. Alhwaiti, N. Alshammari, and M. H. Siddiqi, "Boosting breast cancer detection using convolutional neural network," *Journal of Healthcare Engineering*, vol. 2021, p. 1–11, 2021.

[4] A. A. M. Al-Saffar, H. Tao, and M. A. Talab, "Review of deep convolution neural network in image classification," *2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, 2017.

[5] A. I. Pritom, M. A. R. Munshi, S. A. Sabab, and S. Shihab, "Predicting breast cancer recurrence using effective classification and feature selection technique," in *2016 19th International Conference on Computer and Information Technology (ICCIT)*, pp. 310–314, 2016.

[6] N. Zhang, Y.-X. Cai, Y.-Y. Wang, Y.-T. Tian, X.-L. Wang, and B. Badami, "Skin cancer diagnosis based on optimized convolutional neural network," *Artificial Intelligence in Medicine*, vol. 102, p. 101756, 2020.

[7] J. A. Jaleel, S. Salim, and R. B. Aswin, "Computer aided detection of skin cancer," *2013 International Conference on Circuits, Power and Computing Technologies (ICCPCT)*, 2013.