Convolution Neural Network Using Deep Learning for Breast Cancer Analysis

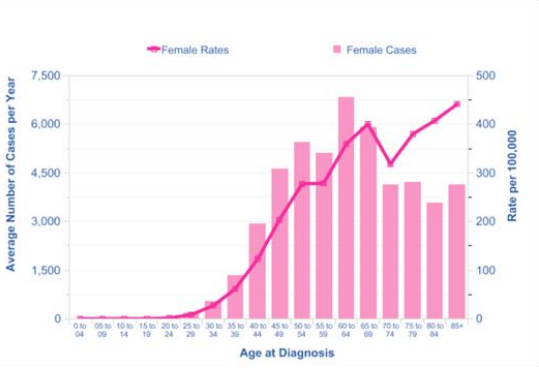
MADA DILEEP KUMAR[1], DEVANA MAHESH BABU[2]
[1, 2] *Computer Science Engineering, Hindustan Institute of Technology and Science, Chennai, India*

Abstract- ***Breast Cancer is one of the major cancer type that is causing 2nd highest fatality rate in women all over the world. It is detected by various technical methods like MRI scans, City scans etc., and combining the results with manual identifications like lumps at breast nodes etc., But it is very difficult to detect cancer accurately by the doctor, surgeon, or a pathologist. It requires years of experience to predict better results manually. To overcome this problem, Doctors started depending on most powerful techniques like deep learning and artificial intelligence. The images of various cases were taken and are analyzed using various methods. Convolution Neural Networks which is mainly used in image side works are used to classify and analyze the scanned images. In recent years, Histopathology is used in medical sciences abundantly by which better results are getting produced. Pathology is a science which mainly works with effects and causes of various tissues. The images of those tissues are called whole-slide images. They are generally very big in size and occupy more amount of memory. A CNN model was developed to overcome this problem case where a better accuracy is produced with respect to the other models which also worked on the same objective.***

Indexed Terms- ***Breast Cancer, Histo-pathology images, Microscopy Whole-slided images.***

I. INTRODUCTION

Breast cancer is one of the leading causes of cancer-related death in women around the world. Over 26, 000 women were diagnosed with breast cancer in Canada in 2017 which represents 25% of all new cancer cases in women. By the Year, more than 5, 000 women in Canada lost their lives due to breast cancer. It is evident that early diagnosis can significantly increase treatment success. Many methods are emerging to handle and treat the patients suffering from various types of breast cancers. Increase in the technology has made a bigger leap in the development of infrastructure in the field of medical science. Diagnosis in breast cancer is increasing thoroughly every year. It is becoming difficult for the manual interference alone to handle the ever growing cases world-wide. The main victims in the most cases are women. Various reasons like lack of awareness about the disease, complete dependency on the manual results, ignorance were the result of this problematic situation. The surgeons and pathologists sometimes make mistake in analyzing the symptoms of the patient and wrongly predict the results. It requires years of experience to correctly analyze whether the problem of the patient is actually a cancer or a normal lump node. This increased the substantial growth of the cases every year.



Graph.1 Annual statistics of Breast Cancer

The above graph shows how age at diagnosis of patients vs average number of cases per year increasing world. We can observe how steep the graph is increasing over time on Australia.

A. Related work

According to the number of diagnostic classes and image types (Microscopy and whole slide images (WSI)), an intelligent system is developed for early

detection and treatment of breast cancer which benefits our societies.

B. Histology and whole slide images (WSI)

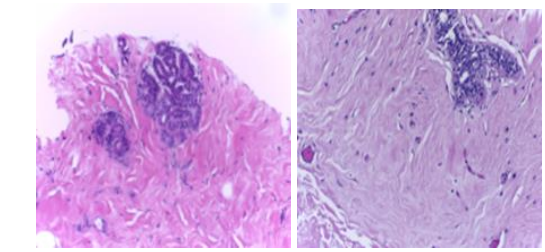
The dataset consists of high resolution images that are used and are used to extract the relevant features from them imposing extra limitations to implementing a feed forword convolution network. To train a CNN on high resolutional images requires more memory, which is not available in most of the scenarios. Thus, these images are divided into batches and they are tranfered into conolutional layers to find the which type of breast cancer is found in the image.

Our dataset is composed of 400 high resolution breast histo-pathology microscopy images. The dataset was available at https://iciar2018-challenge.grand-challenge.org/dataset/.

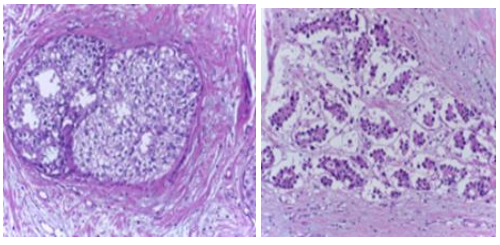
The dataset is composed of Haematoxylin and eosin (H&E) stained breast histology microscopy and whole-slide images.

Microscopy images are ecompu as normal, benign, insitu carcinoma or invasive carcinoma according to the predominant cancer type in each image.

The distribution of Images in the dataset is as follows Normal, Benign, In-Situ and Invasive.



a. Normal b. Benign



c. In-situ d. Invasive

Fig.1 Sample Images in dataset

Whole-slide images are high resolution images containing the entire sampled tissue. In this sense, microscopy images are just details of the whole-slide images. Because of that, each whole-slide image can have multiple normal, benign, insitu carcinoma and invasive carcinoma regions. The annotation of the whole-slide images was performed by two medical experts and images where there was disagreement were discarded. Each image has a corresponding list of 17ecompu coordinates that enclose benign, in situ carcinoma and invasivec arcinoma regions (the remaining tissue is considered normal and thus is not relevant for performance evaluation).

II. ANALYSIS MODEL FOR BREAST CANCER

These convolution networks are more efficient when they are trained reliably.

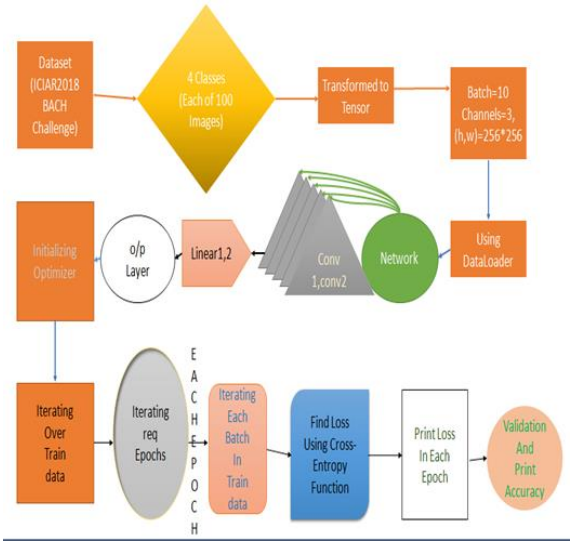


Fig.2. Design and Dataflow

As shown in the above figure, the, model we proposed is a. simple CNN model which has 5 layers in it. They have 2 convolution layers, 2 linear layers and one output layer. We took the input size of the images 2048*1536.Then we resized them to, 256*256 size for convenience with our network. Firstly, took the dataset from ICIAR2018 BACH Challenge .It contains four classes .In each calss hundred images are available. And those hundred images each are stored in their respective folders such as Benign, Normal, in-situ and invasive.

In the model, took the images and transform its size into required size and then they are converted into tensors which plays main role in the program. These tensors contain the resized images in it. These tensors are loaded into a data loader which is the capable of fetching the data from the tensors and convert into batches which contains the sample data. This sample data is now having 4 types of data or parameters such as batches, channels. height and width of the sample data. By loading all the data into the data loader it takes the data into the network class. The network class consists of convolution layers and these convolution layers are used to find the sailent features in the images available in the dataloader. Convolution layer consists of in-channels and out-channels which is filters to find the sailent features. After passing through the convolution layers the data is passed through linear layer to get the required output in the existing no of classes. Then the network is initialized using an optimizer which is used to reduce the amount of loss memory. Here, adding the gradient will helping finding the weights of the tensors. This will helps in finding the lowest weights. After this applied the loss fuction called Cross-Entropy to calculate the loss in the tensors. And here model calculate the same model to get the lowest by using iterations to get the decreased loss. At last the validation and testing process is done and print the accuracy and loss to the each epoch.

C. Performance and validation

This above model produced an accuracy of 96% which is better than other models and works done previously. Validation is done using validation data by comparing real data without labels and data with labels. The data is passed through CNN and the result is achieved.

III. DISCUSSION

This model considered the problem of breast cancer classification using microscopy tissue histopathology images. By utilizing deep learning techniques and proposed a CNN network solutions to overcome the hardware incapabilities produced by the usage of very large images. This work mainly contributes presenting a clear idea of how different models with different properties are used which helps to process very large scale images.

Title	Author	Architecture	Images in dataset
Two Stage CNN for Breast Cancer Histology Image Classification	Kamyar Nazeri, Azad Aminpour, and Mehran Ebrahimi	VGG	400
Breast Cancer Classification in Histopathological Images using CNN	juanying xie,Ran Liu,Joseph Luttrell IV,Chaoyang Zhang	InceptionResNet V2	7909
Breast Cancer Classification in Histopathological Images using Convolutional Neural Network	Mohamad Mahmoud Al Rahhal	VGG	7909
Deep Learning Based Analysis of Histopathological Images of Breast Cancer	Abdullah-Al Nahid , Mohamad Ali Mehrabi, and Yinan Kong	RNN	7909
Breast cancer histopathology image classification through	Chuang Zhu1, Fangzhou Song, Ying	Hybrid CNN	7909

assembling multiple compact CNNs	Wang, Huihui Dong, Yao Guo and Jun Liu		
Classification of Histopathological Biopsy Images Using Ensemble of Deep Learning Networks	Sara Hosseinzadeh Kassani, Peyman Hosseinzadeh Kassani, Michal J. Wesolowski, Kevin A. Schneider, Ralph Deters	Vgg,denseNet,Mobilenet	3271680

Table.1 Comparison of various Models

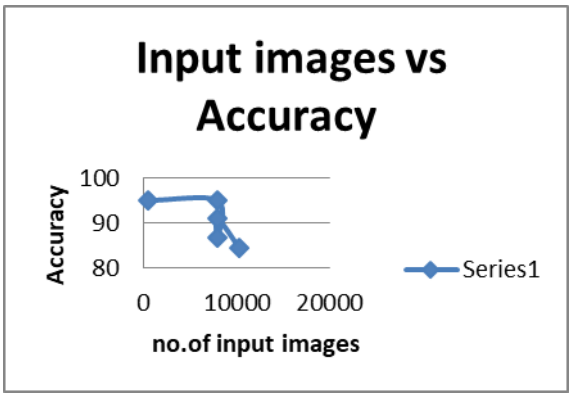
Title	Accuracy	Input Image size	Classes
Two Stage CNN for Breast Cancer Histology Image Classification	95	2048,1536	4
Breast Cancer Classification in Histopathological Images using CNN	95	700,460	2
Breast Cancer Classification in Histopathological Images using Convolutional Neural Network	86.6	700,460	2
Deep Learning Based Analysis of Histopathological Images of Breast Cancer	91	700,460	2
Breast cancer histopathology image classification through assembling multiple compact CNNs	84.4	700,460	4
Classification of Histopathological Biopsy Images Using Ensemble of Deep Learning Networks	98.64	96,96	2

Table. 2 Comaprison of Various Models

Title	Batch size	Optimizer	Learning rate
Two Stage CNN for Breast Cancer Histology Image Classification	64	Adam	0.001
Breast Cancer Classification in Histopathological Images using CNN	32	Adam	0.002
Breast Cancer Classification in Histopathological	50	Mini Batch	0.01

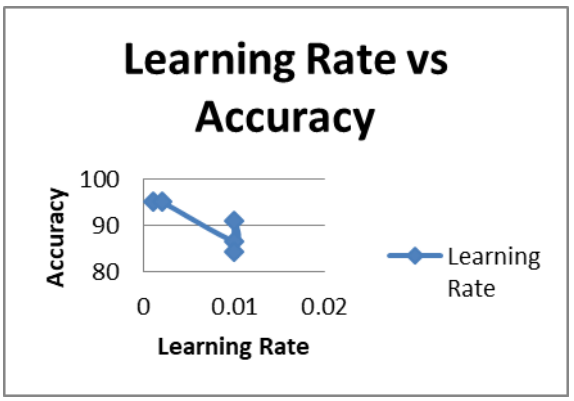
Images using Convolutional Neural Network			
Deep Learning Based Analysis of Histopathological Images of Breast Cancer	64	SGD	0.01
Breast cancer histopathology image classification through assembling multiple compact CNNs	64	SGD	0.01
Classification of Histopathological Biopsy Images Using Ensemble of Deep Learning Networks	32	Adam	0.6,0.8,0.001

Table.3 Comaprison of Various Models



Graph.2 input image vs Accuracy

This graph shows the deviation of accuracy with respect to no of input images.



Graph.3 Learning Rate vs Accuracy

This graph shows the deviation of accuracy with respect to learning rate.

CONCULSION AND FUTURE ENHANCEMENT

The main consideration is one of the crucial problem in Women, Breast cancer. The unstructured data i.e., the histo-pathology images of breast cancer of various patients are taken to classify and analysis into four types of breast cancer Benign, Normal, in-situ and invasive. We used emerging technologies like Deep learning for image recognition using CNNs to classify the problem. Though, the network is small and more epochs are needed, model is getting a better accuracy than other models which is a major advantage. The accuracy produced was nearly 96%. The proposed thesis can be helpful in extending in many ways. Some of them are the design of the network can be altered to produce various different models. The structured data of the patient can also be accompanied with this network model to produce better results. Other image data like MIR scans and CT scans can also be accompanied with data in the proposed model and compare results. Making the system as possible as efficient by dec4reasing the time cost and hardware requirements.

REFERENCES

[1] Two-Stage Convolutional Neural Network for Breast Cancer Histology Image Classification- Kamyar Nazeri, Azad Aminpour, and Mehran Ebrahimi.[1]

[2] Breast Cancer Classification in Histopathological Images using Convolutional Neural Net4work-juanying xie,Ran Liu,Joseph Luttrell IV,Chaoyang Zang.[2]

[3] Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering - Mohamad Mahmoud Al Rahhal .[3]

[4] Deep Learning Based Analysis of Histopathological Images of Breast Cancer-Abdullah-Al Nahid , Mohamad Ali Mehrabi, and Yinan Kong.[4]

[5] Breast cancer histopathology image classification through assembling multiple compact CNNs-Chuang Zhu1, Fangzhou Song, Ying Wang, Huihui Dong, Yao Guo and Jun Liu.[5]

[6] Classification of Histopathological Biopsy Images Using Ensemble of Deep Learning Networks-Sara Hosseinzadeh Kassani, Peyman Hosseinzadeh Kassani, Michal J. Wesolowski, Kevin A. Schneider, Ralph Deters.[6]