BREAST CANCER DETECTION USING KERNEL NEURAL NETWORK

Geetha.M[1] Khyathy.V[2] Pokala Hemavathi[3]

[1]Assistant Professor, Department of CSE, Panimalar Institute of Technology, Chennai

*[2],[3]*UG Scholar, Department of CSE, Panimalar Institute of Technology, Chennai

Abstract - *Among the assorted treatment strategies obtainable for carcinoma ,Histopathological Classification may be a normal for diagnosis of cancer. Carcinoma classification is based totally on pictures of the tissue within the tumour. We tend to classify the pictures of tissue into 2 major classes of tumour : Benign and Malignant exploitation KNN . Existing works use CNN , we tend to use KNN over CNN since it provides indefinite complicated actions of the human recognition system by creating use of the kernel trick. KNN may be a generalized version of convolution which may enhance the model's extent and might capture higher order of traits exploitation of reinforcement kernel functions***.**

Keywords - BreastCancer, Kervolution, Classification, Histopathology.

1. INTRODUCTION

Deep Learning is an emerging technique which has caught the interest of various scientists and researchers. Deep learning a variant in machine learning technology which lets machines learn from training, skill, data, futures, and helps computers comprehend a collection of concepts and rules for the problem of the real world. The need of the human machine operator to determine all the details a machine needs is depreciating as device collects the information or data from expertise. The architectures like CNN, SVM, KNN helps the machine to learn features hidden in the data collection which in turn helps the networks for classifying the data into groups. Multi-layered perceptron is a nice example of deep learning where layers are fully connected. In this project paper, we used kervolutional neural network for classifying breast cancer images. In kervolutional neural networks, the learning network architecture follows the same method as that of Convolution neural networks except for the convolution layer we use Kervolution layer. Breast Cancer has struck around 2.1 million women worldwide and additionally a reason for the biggest number of cancer-related death among ladies. Breast cancer is one out of the many most general diseases among women. In 2018, 627,000 people were reported to be dead with breast cancer–about 15% of the total deaths from cancer. Although women in more developing countries are at higher rates of breast cancer, rates are increasing worldwide in nearly all areas. To improve outcomes of breast cancer and for survival, early detection of the type of cancer is important, with advancements in technologies and with the availability of a large amount of collection of previous patient's data has motivated many researchers and scientists to help the pathologists in diagnosing Histopathological Images.

2.DATASET

The BreakHis data set contains minuscule level biopsy pictures of both malignant and benign breast cancer tumors. It is a set which consists of 9109 micro-sized pictures of breast growth tissue gathered from around eighty-two patient's treatment with totally four diverse magnifying factors.40X, .100X,.200X and.400X. In which 2480 are benign whereas 5429 are malignant sample pictures and the info was designed with the assistance of P and D science lab, Brazil. Table 1 shows the information sets structure belonging to two classes benign and malignant with totally 4 different magnification factors. The data set contains several types built on the way the cells look from underneath the microscope which can have different prediction and treatment results. The data set carries 4 histological discrete types belonging to benign breast tumors and 4 malignant type tumors. The dataset distribution is represented in Table 1.

Table 1: Image Data samples.

Magnification	Benign	Malignant	Total
40X	652	1370	1995
100X	644	1437	2081
200X	623	1390	2013
400X	588	1323	1820
Total Images	2480	5429	7909

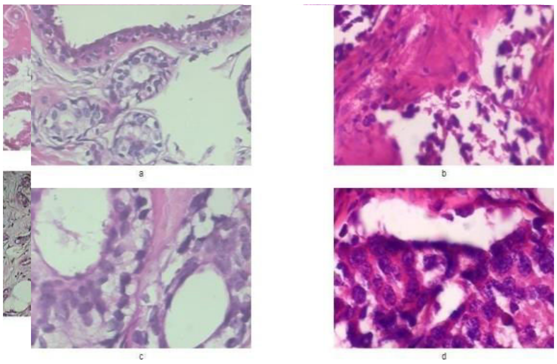


Figure 1.40X(a,b)-100X(c,d)

Figure 2.200X(a,b)-400X(c,d)

Figure 1 and Figure 2 shows sample images belonging to 2 classes Benign and Malignant. Benign figures are Figure 1-a, Figure 1-c, Figure 2-a and Figure 2-c. Malignant figures are Figure 1-b, Figure 1-d. Figure 2-b, Figure 2-d of 40X magnification, 100X magnification, 200X magnification and 400X magnification respectively.

3.RELATED WORK

Many researchers and scholars have worked and have presented their work on breast cancer prediction and classification tasks. Observed that many of the related works were done using Deep learning methodologies and were done using Convolution Neural Networks. In this work,Kervolutional Neural Network is used to classify BreakHis data set which was modified for future work, Kervolutional Neural Network uses the kernel tricks by using a more generalized version of Convolution which also intensifies the model's extent and can capture higher order of traits using reinforcement kernel functions from images.

4 .KERVOLUTIONAL NEURAL NETWORK

Most of the computer vision tasks like image recognition, object identification have been seen to be successful by using convolutional neural networks whose core function convolution was inspired by the animal visual cortex which is known as the receptive field. It leverages upon its equivariance for its translations for improved performance of machine learning or deep learning model. Convolution is optimal for linear classifiers but for non- linear classifiers, the convolution needs many parameters and the complexity increases and even with introduction of kernel tricks in convolution, it cannot extract non- linear patch wise attributes. To solve problems such as these, Kervolution is used. Starting with a convolution using a vectorized input and continuing with a non-linear mapping function which allows us to extract attributes from a higher dimension space and using the kernel trick for bypassing explicit calculation of higher-order features. Kervolution uses the advantages of convolutional neural networks with the addition of new features like sharing weights, equivariance to translation and increase in model capacity.

In this project , the BreakHis data set is rearranged removing all the subtype cancer from the data set and converting that data set into only 2 types of cancer benign or malignant based on their magnification. The model architecture used was the simple deep neural network replacing Convolution layer with Kervolution layer.The model is trained for 50 epochs for each type of magnification images separately with the same architecture andyper parameters.Figure 3 displays the block diagram for the classification task where the rearranged data set is split in ratio 8:1:1 for training, validating and testing. At first, the model is pinned using the feature extracted from training data and the model is analyzed after each epoch using the validation dataset and finally that trained model is

The neural network consists of a total of 4 blocks

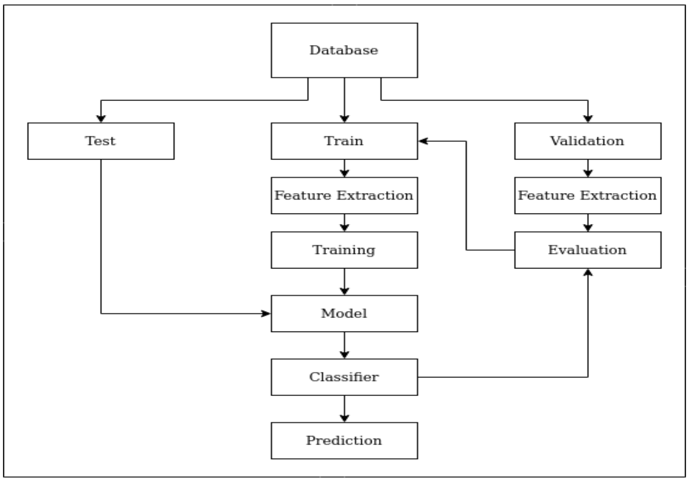


Fig 3. Block diagram for the Classification Process

The data set under the Testing phase is directed directly to the Classifier state and then from the Classifier state the data is classified to 1 of the 2 recognized classes. The data set under the Train phase is given to the Feature Extraction state and after the features are extracted from the data, the extracted features data is given for training and the output received during the training phase is given to the

5. METHODOLOGY

5.1ARCHITECTURE

where the 4th block is a hidden block. The 1st block consists of the Input layer, Kernel Convolution layer, Activation-Relu, Batch Normalization, Dropout. Input image was of size 48*48*3 and the same was used for the batch normalization layer. Pooling layer was followed by dropout with drop rate of 0.25 and a dimension of 24*24*32. then again followed by Kernel Convolution and relu activation. 2nd block with 2 repetitions of kervolution layer with our Relu activation and max-pooling of weights (2,2) and (12,12,64) is done after performing batch normalization and dropout layer with drop rate 0.25 and dimensions of (12,12,64). 3rd block with 3 repetitions of Kervolution layer with relu activation and max-pooling of (2,2) and (6,6,128) is done after performing batch normalization and drop out layer with drop rate 0.25 and with dimensions of (6,6,128). 4th and the final block consists of 1 flattening layer and dense layer with 256 units and relu activation layer continued with batch normalization and dropout layer with 0.5% drop rate followed by another dense layer and finally SoftMax activation layer for classification . Table 2 below shows the neural network architecture used with their respective parameters .

Table 2: Deep Neural Network Architecture

model for getting the result. Once the result is obtained it is then directed to the evaluation state which is then compared with the data set stored in the validation phase and the values are noted down.

Block	Operation	Dimensions
	Input	(48,48,32)
	Kernel Convolution	(48,48,32)
	Activation-Relu	(48,48,32)
	Batch Normalisation	(48,48,32)
	Max Pooling	(2,2),(24,24,32)

1	Drop out	(0.25),(24,24,32)
	Kernel convolution	(24,24,64)
	Activation-Relu	(24,24,64)
	Batch Normalisation	(24,24,64)
	Kernel Convolution	(24,24,64)
	Activation -Relu	(24,24,64)
	Batch Normalisation	(24,24,64)
2	Max Pooling	(2,2),(12,12,64)
	Dropout	(0.25),(12,12,64)
	Kernel Convolution	(12,12,128)
	Activation-Relu	(12,12,128)
	Batch Normalisation	(12,12,128)
	Kernel Convolution	(12,12,128)
	Activation-Relu	(12,12,128)
3	Batch Normalisation	(12,12,128)
	kernel Convolution	(12,12,128)
	Activation-Relu	(12,12,128)
	Batch Normalisation	(12,12,128)
	Max Pooling	(2,2),(6,6,128)
	Dropout	(0.25),(6,6,128)
	Flatten	4608
	Dense	256
	Activation-Relu	256
4	Batch Normalisation	256
	Dropout	(0.5),256
	Dense	256
	Activation-SoftMax	2

5.2 LOSSES AND OPTIMIZERS

The binary cross-entropy for loss and adagrad with a learning rate of 1$e-2$ and decay rate with learning rate/number of epochs (□$e-2/50$) for optimizing .

6.RESULTS

The Break his data set is divided into 3 parts of which 80% for training and 10% is for validation and 10% for testing the model. The data set is divided in such a format that there is separate data for testing the trained model so that our model can support classifying the new data given. This model was trained on CPU using tensor flow framework. System specs Intel® CoreTM i5-8250U CPU @1.60GHz 1.80GHz with 8GB ram with 64-bit Windows operating system.

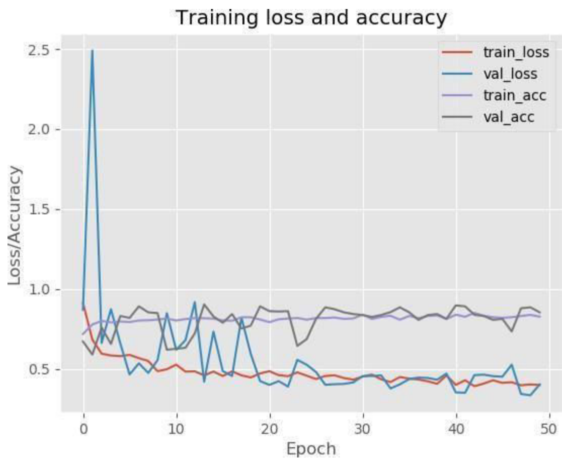
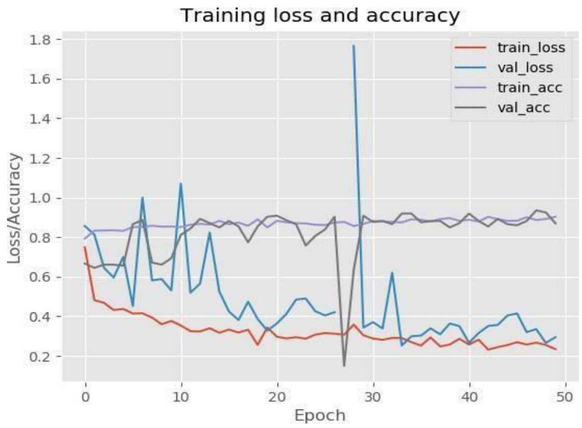


Figure 4. 40X Magnification

Figure 4 shown above depicts the training loss and training accuracy in addition to the validation accuracy and validation loss for 40X magnification images for the respective number of epochs.

Figure 5. 100X Magnification

Figure 5 shown above depicts the training loss and



training accuracy in addition to the validation accuracy and validation loss of 100X magnification for the respective number of epochs it is Benign.

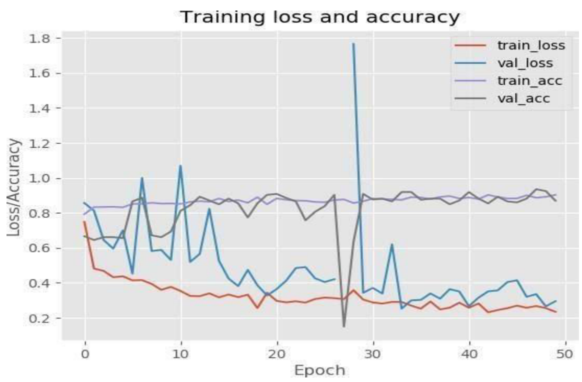


Figure 6. 200X Magnification

Figure 6 depicts the training loss and training accuracy in addition to the validation accuracy and validation loss of 200X magnification for the respective number of epochs.

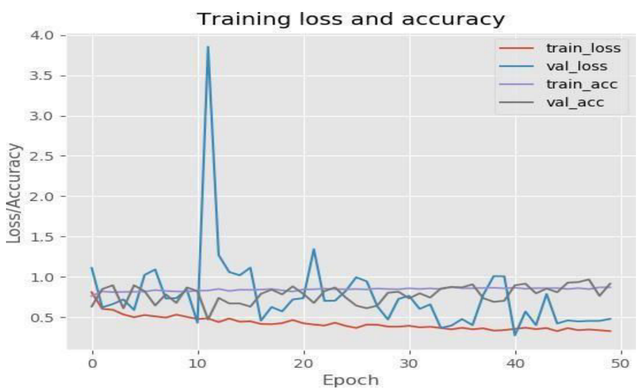


Figure 7. 400X Magnification

Figure 7 shown above depicts the training loss and training accuracy in addition to the validation accuracy and validation loss of 400X magnification for the respective number of epochs.

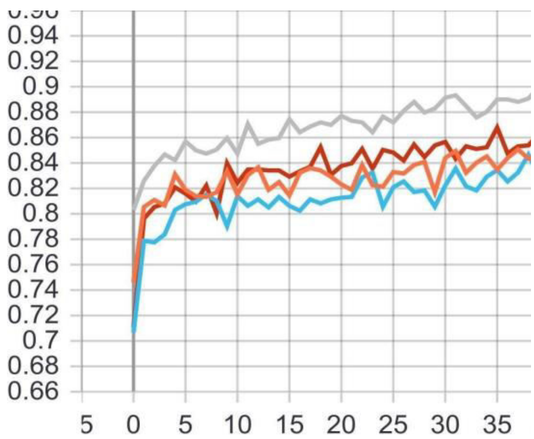


Figure 8. Model Training Accuracy

Figure 8 show the graph regarding the training accuracy of the model when it was executed considering the 4 magnifications.Here, the Blue color line points to 40X magnification, the Orange color line points to 100X magnification, the Grey color line points to 200X magnification, the Red color line points to 400X magnification.

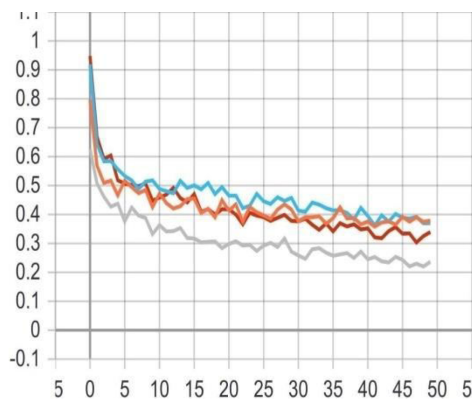


Figure 9 shows the graph regarding the training loss of the model when it was executed considering the 4 magnifications. Here, the Blue color line points to

40X magnification, the Orange color color line points to 400X magnification line points to 100X magnification, the Grey color line points to 200X magnification, and the Red.

Figure 9 .Model Training loss

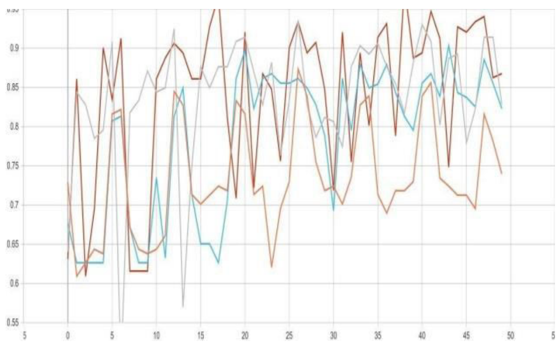


Figure 10. Models Validation Accuracy

Figure 10 shows the graph regarding the validation accuracy of the model when it was executed considering the 4 magnifications. Here, the Blue color line points to 40X magnification, the Orange color line points to 100X magnification, the Grey color line points to 200X magnification, the Red color line points to 400X magnification.

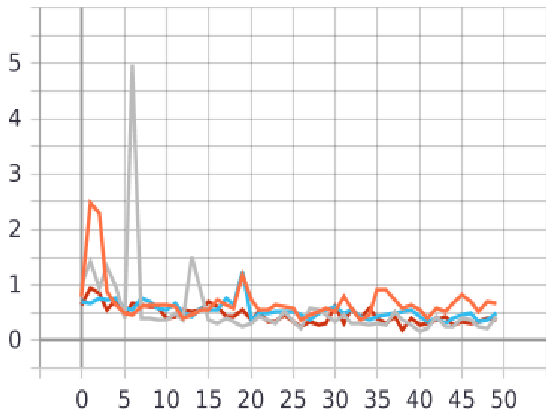


Figure 11 . Models Validation Loss

Figure 11 displayed above shows the graph regarding the validation loss of the model when it was executed considering the 4 magnifications. Here, the Blue color line points to 40X magnification, the Orange color line points to 100X magnification, the Grey color line points to 200X magnification, the Red color line points to 400x Magnification.

Table 3 : Classification report for different image magnification

Magnification (X)	Class	Precision	Recall	F1 Score	Support
40	Benign	0.95	0.68	0.76	63
	Malignant	0.86	0.99	0.92	138
100	Benign	0.70	0.32	0.44	65
	Malignant	0.75	0.94	0.84	144
200	Benign	0.82	0.82	0.82	62
	Malignant	0.93	0.93	0.9	155
400	Benign	0.79	0.78	0.79	59
	Malignant	0.90	0.90	0.90	123

Table 3 shows the best classification reports observed during the work.

Table 4 : Kervolutional neural network models efficiency for different image magnification.

Magnification(X)	Accuracy	Sensitivity	Specificity	Support
40	0.8756	0.6825	0.9855	201
100	0.7464	0.3230	0.9375	209
200	0.8986	0.8225	0.9290	217
400	0.8626	0.7796	0.9024	182

Table 4 shows the model's accuracy, sensitivity and specificity obtained while testing our trained model with support as the number of images used while testing the model.

7.EXISTING SYSTEM

The existing system works well with the 40X,200X and 400X and lacks the stability with the 100X magnification of the reports .The accuracy of 40X , 200X and 400X is quite stable when it is compared with the 100X magnified images or data report .

8. CONCLUSION

kervolution neural networks used for converting the convolution layer with kervolution layer with some changes in the parameters and tested it on BreakHis data set for histopathological images using tensor flow for classifying images to Benign and Malignant tumors. In this work, it is seen that the model performing better when compared to models that are using Convolutions for 40X, 200X and 400X magnification factor images still it is seen that there are more options for improvement as currently, the model could not give required accuracy for 100X magnification, and can see that during validation the model is not quite stable enough for giving linear increments in accuracy or decrements in the loss.

9.FUTURE ENHANCEMENT

Future works include designing a more stable network model and increasing the model's overall accuracy and providing more acceptable evidence to prove kervolution provides more capacity to the deep learning models in learning more features compared to convolution.

10.REFERENCES

[1] Szegedy Y, Jia, Sermanet W, Liu P, Reed S, Anguelov D, Rabinovich A (2015) Wang J, Yang L, Xie J, Yuan (2019) Kervolutional Neural Networks. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp31–40.

[2] Espanhol FA, Oliveira LS, Petitjean C, Heutte L (2016b) A Dataset for Breast Cancer Histopathological Image Classification. IEEE TransactionsonBiomedicalEngineering63(7):1455–1462,DOI10.1109/tbme.2015.2496264,

[3] Kumar K, Rao ACS (2018) Breast cancer classification of image using convolutional neural network. In: 4th International Conference on Recent Advances in Information Technology (RAIT), pp1–6.

[4] Arevalo J, González FA, Ramos-Pollan R, Oliveira JL, Lopez MAG (2015) Convolutional neural networks for mammography mass lesion classification. 37th Annual International Conference of the IEEE pp 797–800.

[5] Bernhard E, Boser IM, Guyon VN, Vapnik (1992) A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory pp 144–152.

[6] Blondel M, Ishihara M, Fujino A, Ueda N (2016) Polynomial Networks and Factorization Machines: New Insights and Efficient Training

Algorithms. International Conference on MachineLearning.

[7] Bo L, Sminchisescu C (2009) Efficient Match Kernel between Sets of Features for Visual Recognition. Conference on Neural Information Processing Systems pp 135–14

[8] Bo L, Ren X, Fox D (2010) Kernel Descriptors for Visual Recognition. Conference on Neural Information ProcessingSystems.

[9] Cohen N, Sharir O, Shashua A (2016) Deep SimNets. IEEE Conference on Computer Vision and Pattern Recognition pp4782–4791.

[10] Cui Y, Zhou F, Wang J, Liu X (2017) Kernel Pooling for Convolutional NeuralNetworks.

[11]Dalal N, Triggs B (2005) Histograms of Oriented Gradients for Human Detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition pp 886–893.

[12] Dubey A, Gupta O, Guo P, Raskar R, Farrell R, Naik N (2018) Pairwise confusion for fine grained visual classification. Proceedings of the European Conference on Computer Vision (ECCV) pp70–86.

[13]Hafemann G, Oliveira LS, Cavalin P (2014) Forest species recognition using deep convolutional neural networks. Pattern Recognition (ICPR) 2014 22nd International Conference on pp 1103–1107.

[14]Jia Deng W, Richardsocher, Li-Jiali, Kaili, Fei-Fei (2009) Imagenet: A large- scale hierarchical image database. Computer Vision and Pattern Recognition pp 248–255.

[15]Jiao Z, Gao X, Wang Y, Li J (2016) A deep feature-based framework for breast masses classification. Neurocomputing 197:221–231.