OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

DETECTION OF BREAST CANCER USING DEEP NEURAL NETWORK TECHNIQUES

Prajakta Khendad[1], Dr. B. K. Sarkar[2]

Department of Computer Engineering Padmabhooshan Vasant dada Patil Institute of Technology, Bavdhan Pune SavitribaiPhule, Pune University , Maharashtra[1,2]

prajaktakhendad@gmail.com,dr.bksarkar2003@gmail.com

Abstract: In worldwide woman's death rate increases due to the breast cancer. That's why medical imaging techniques are used to detection and diagnosis of breast cancer. Automated classifiers could significantly modernize the diagnostic process in terms of both accuracy and time requirements by automatically distinguishing benign and malignant tumors. In this paper propose a new breast cancer diagnosis technique such as DNN (Deep Neural Networks) to the prognosis prediction of women's breast cancer. Here, NN (Neural Networks) plays an important role in this especially in the application of breast cancer detection.

Keywords: Brest Cancer, Benign and Malignant Tumors, Neural Network, Deep Neural network and Images.

I INTRODUCTION

One of the main current challenges and opportunities in biology is the interpretation of the growing volume of single-cell data with high bandwidth and large amount of information. Here, consider 46 visualization data of fluorescence microscopy [1], in particular cytometry imaging flow 47 [2]. IFC (Imaging Flow Cytometry) it combines 48 high-throughput capabilities of flow cytometry and fluorescence sensitivity with thesingle-cell images. The corresponding 49 fluorescent labels are selected to evaluate certain phenotypes. The large number of 50 single cells analyzed per sample-often hundreds of thousands-makes the Imaging 51 flow cytometry extraordinarily suitable for deep learning that requires very large 52 training sets.

Breast cancer is one of the causes of death in women, most often diagnosed without skin cancer in women. Breast cancer occurs when the mammary tissue becomes abnormal and uncontrollable to divide. These abnormal cells form large chunks of tissue, which later turn into tumors. Such disorders can be successfully treated if they are detected at an early stage. Therefore, it is important to have a proper screening method for the earliest signs of breast cancer. Automated classifiers help radiologists distinguish between benign and malignant tumors. Here, ANN (Artificial Neural Network) which functions as an automatic classifier was examined. In the field of medical image processing, ANN has been applied to various data classification and pattern recognition tasks and has become a promising classification tool for breast cancer [4].

In the early 1980s, there was an increase in the use of neural networks in image and signal processing. The main advantage is the reduction of manipulation time due to parallel-distributed processing of neural networks [6]. The network was then widely used in common image processing techniques such as vector quantization, eigenvector extraction, two-dimensional pulse code modulation, or two-dimensional filtering [7].An artificial neural network resembles a function of a biological neuron, and it consists of neurons with different layers, and these neurons are linked by numerical weights; these weights can be changed due to the learning behavior of the network to get closer to the optimal result. Usually, in image processing applications, the number of neurons is directly related to the number of pixels in the input image [8], and the number of layers depends on the processing steps.

In this paper study about the related work done, in section II, the proposed approach modules description, mathematical modeling, algorithm and experimental setup

in section III .and at final provide a conclusion in section IV.

II LITERATURE REVIEW

Here discussed the literature review of existing techniques:

In automatic detection of the cell compartment containing the fluorescent labeled protein problems and it is a relatively simple task for the experienced person, but difficult to automate on a computer. Here they teach the 11-layer neural network according to the mapping of thousands of yeast proteins, reaching a classification accuracy of localization in cell 91% and accuracy of protein 99% of the held images. Here, they confirm that the low-level network features correspond to the basic characteristics of the image, while the deeper layers separate the localization classes. Using this network as a functional calculator, they train standard classifiers that assign proteins to previously invisible compartments after observing only a small number of training examples. [1].

Deep flow for large data sets of cell cycling Jurkat cells. To begin with, the cell continuation progress by the cell cycle was reconstructed from the raw image data. In this, theyshow that Deep-Flow can train a continuous distance measure between category phenotypes. Second, they can detect and isolate dead cell populations, while the data set had been cleaned using established approaches. Deep-Flow detects these morphologically anomalous subpopulations in an unsupervised fashion. Third, in the label-free classification of cell cycle phases, they achieved a 6-fold reduction in error rates, compared with the recent approach based on a boost to a series of imaging features [2].

In this paper, they propose neural network architecture for the classification and segmentation of microscopyimages within cell populations. This method isbased on the similarity between the aggregation function used in the MIL and the grouping layers used in CNN's. To facilitate aggregation through a large number of instances on CNN maps, they present the pooling and noisy function. The combination of CNN's with thousands to train CNN's using microscopy images of a whole with level labels of image. They demonstrate that the formation of end-to-end thousand CNN's outperforms several previous methods, both in mammalian and data sets of the yeast, without the need of any step of segmentation[3].

In this work [4], they proposed Convolutional Neural Networks (CNN) function for counting the number of colonies contained in confluent agglomerates, which obtained an overall accuracy of 92.8% in a large a challenging dataset. The proposed CNN-based technique to estimate the cardinality of colony aggregates exceeds traditional imaging approaches, becoming a promising approach for many related applications.

In this paper [5], they used CNN (Convolutional Neural Networks) with deep max-pooling to detect mitosis in images of breast histology. Networks are able to classify each pixel in the images, using a Pixel-centered patch as a context. Simple post-processing then applies to network output. This approach won the ICPR 2012 mitosis detection competition. The different image areas are characterized by different types of fabric, which display a very variable appearance.

The backpropagation technique is used for the learning of the multilayer neural network. It is most commonly used as a learning algorithm for neural networks to handle continuous data and differentiable functions for both single and multi-layer models. Slow convergence and lack of training time is a disadvantage, in that case, the traditional BA propagation algorithm compared to other sports methods. In recent years there has been a lot of improvement and modification of the learning algorithm, to overcome the problem. In this paper the propagation algorithm of the several variations of the algorithm is proposed for improved functionality is thoroughly examined [6].

This is necessary to efficiently map the structure and 3D connectivity of the brain. The segmentation of biological fat membrane uses an artificial deep neural network as a Pixel classifier. The label for each pixel (membrane or non-membrane) is predicted from the raw pixel values in a square window centered on it. The input layer maps each of this pixel in the window to FAS ton. It is followed by a succession of convolutional and Max-pooling layers that preserve 2D information and extract features with increasing levels of abstraction [7].

The system automatically detects segments and locates cells and nuclei in microscopic images. The system was designed as the central component of a fully automated phenotyping system. The system contains three modules (1) a Convolutional network trained to classify each pixel into five categories: cell wall, cytoplasm, core membrane, core, outer Medium; (2) an energy base Model that cleans the exit of the Convolutional network through local learning consistency

restrictions to be satisfied by the images on the label; (3) a set of elastic embryo models in various stages of development that adapt to the tag images [8].

The author proposesmRMR feature selection method with Multiple Kernel Learning classification method, for the prediction of GBM prognosis. The survival rate of patients is deferent for the every subtypes of glioma. For this experiment they used the cancer genome atlas (TCGA) dataset of various types of cancers. In this they improve the prognosis prediction accuracy of GBM and they compare performance with the one kernel method using same dataset. Because
of the deferent datasets and process method can't compare with other researches directly [9].

III PROPOSED APPROACH

Problem Statement

Implement techniques such as deep neural networks techniques for the predicting prognosis of women breast cancer.

Proposed System Overview

Several DL (Deep Learning) techniques have been proposed in the literature, including Deep Neural Network (DNN), Convolutional Neural Network (CNN), RNN, Deep Auto-Encoder (DA), Deep Boltzmann Machine (DBM), Deep Belief Network (DBN), Deep Residual Network (DRN), and Deep Convolutional Inverse Network Graphics (DCING) and so on. For the sake of brevity, only those that are widely used with biological data are summarized below. However, interested readers are directed to the references mentioned in each section for specific mathematical details behind each of architecture.Fig.1 shows that the proposed system architecture. Here use multimodal deep neural networks by integrating multi-dimensional data techniques for the prognosis prediction of breast cancers. This probabilistic model is based on the Bayesian Networks (BN) for the prognosis of lymph node-negative breast cancer by integrating clinical and genomic data.

In this system architecture has the following modules:

I. Multidimensional Dataset:

Input dataset such as microarray data is high-dimensional which is consists of approximately 2500 gene per patients. The dataset is extracted from the 1980valid breast cancer patient's data of METABERIC trial, which contains multidimensional data among breast cancer such as gene expression profile, CNA profile and clinical profile.

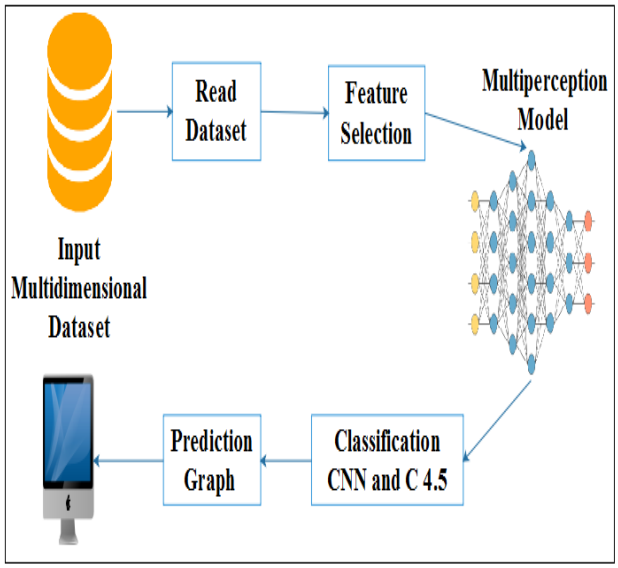

Figure.1.Proposed System Architecture

II. Feature Selections:

Here used mRMR feature selection techniques. It selects the features form the original dataset and it reduces the dimensionality of dataset without a significant loss of features. Features are considered such as nucleus size, cytoplasm size, nucleus grey level and cytoplasm grey level.

III. Multi-Perceptron Mode:

A deep neural network is using for the prediction of the prognosis of women breast cancer. The DNN architecture builds a hierarchy from the hidden layers. Higher level features are extracted implicitly by the combining of lower level features from each layer. A DNN model is composed of an input layer, multiple hidden layers, and output layers. Units between layers are fully connected.

IV. Classification:

SVM, KNN and C 4.5 algorithms are using for the classification. For every round, nine of ten subsets are further divided into the training and validation set, while remaining one subset is for testing. In this way obtain prediction scores of each testing subset after ten rounds and then merge them as an overall prediction score.

V. Prediction Result:

Decision level multimodal fusion is used to integrate both clinical information and breast cancer-specific relationship between genes.

Algorithm

- **Algorithm: C4.5 Algorithm**

Process:

1. Check for the below base cases:
   i. All the samples in the list belong to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class.
   ii. None of the features provide any information gain. In this case, C4.5 creates a decision node higher up the tree using the expected value of the class.
   iii. An instance of previously-unseen class encountered. Again, C4.5 creates a decision node higher up the tree using the expected value.
2. For each attribute a, find the normalized information gain ratio from splitting on a.
3. Let a_best be the attribute with the highest normalized information gain.
4. Create a decision node that splits on a_best.
5. Recur on the sublists obtained by splitting on a_best, and add those nodes as children of the node.

Mathematical Model

Let S be the system such that

S= {I, P, O, SC, FC}

Where,

I=Input of system

O=Output of system

P=Process in the System

SC= Success Conditions

FC = Failure Conditions

I = {Img1; Img2, ….,Imgn};

Process:

1. P1 = {I};
Read dataset Brekhis which contains image of breast cancer.

2. P2 = {P1};
Color features exaction method used to get image information which can be used for classification and CNN.
   a. Color-Based Extraction
   b. Texture-Based Extraction

3. P3 = {P2};
Generation of training and testing files is important. It is generated by taking features extraction from two images.

4. P4 = {P3};
The classification algorithm is used for the classification of breast cancer image by using:
- C4.5
- CNN
- KNN
- SVM

5. P5 = {P4};
A deep learning concept is convolution neural network concept is used for classification and prediction of breast cancer.
Sc = {When successfully detected class of cancer stage}
Fc = {When it fail to detect the class of Breast cancers}

IV. RESULTS AND DISCUSSION

A. *Experimental Setup*

Hardware and software of proposed system given below:

- **Software Technology:**
  1. Technology: Core Java
  2. Tools: JDK 1.8, Netbeans 8.0.2
  3. Operating System: Windows 7
- Hardware Technology
  1. Processor: 1.0 GHz
  2. RAM: 1 GB
  3. Hard Disk: 730 GB

B. *Dataset*

Here, BreakHis dataset used and itcontains microscopic biopsy images of malignant and benign breast tumors. In this BreakHis contains about 8 000 microscopic images of breast tumor tissue images which are collected from the 82 and only four features of images considered such as nucleus size, cytoplasm size, nucleus grey level, and cytoplasm grey level used.

C. *Expected Result*

Table 1 shows, the accuracy comparison between the existing and proposed system algorithm.

Figure 2 shows, accuracy graph between the CNN, C 4.5, KNN, and SVM algorithms. From the graphs, it is concluded that the proposed CNN algorithm is more accurate to predicting prognosis of women breast cancer than C 4.5, KNN, and SVM algorithms.

Table 1: Accuracy Comparison Graph

Algorithm	Accuracy in %
CNN	99.41
C4.5	98.0
KNN	95.44
SVM	78.48

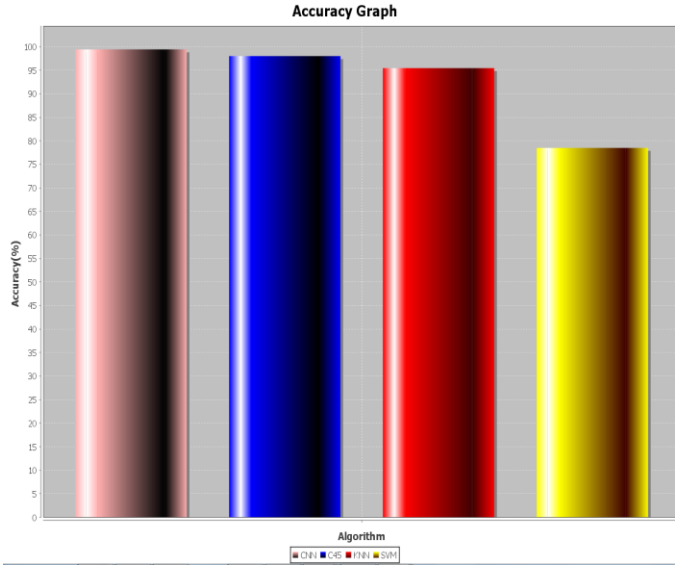

Figure 2. Accuracy Comparison Graph

The figure 3 shows the Precision graph of CNN, C 4.5, KNN and SVM algorithms. The graph shows the proposed system is better than Existing system.

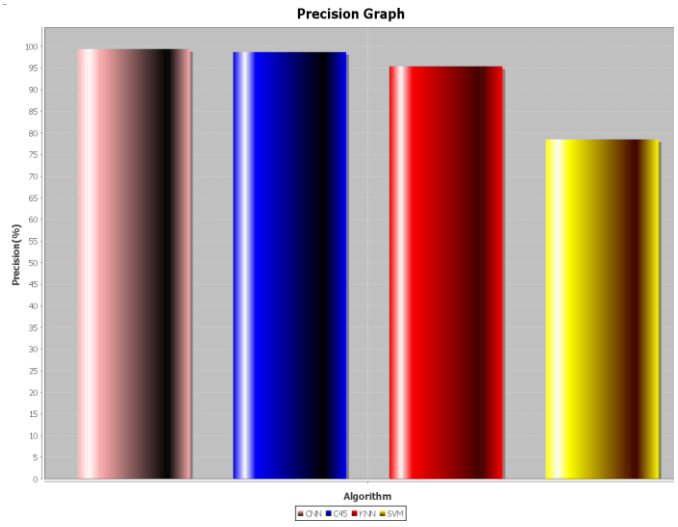

Figure.3. Precision Graph

The figure 4 shows the ROC graph of CNN, C 4.5, KNN and SVM algorithms. The graph shows the proposed system is better than Existing system.

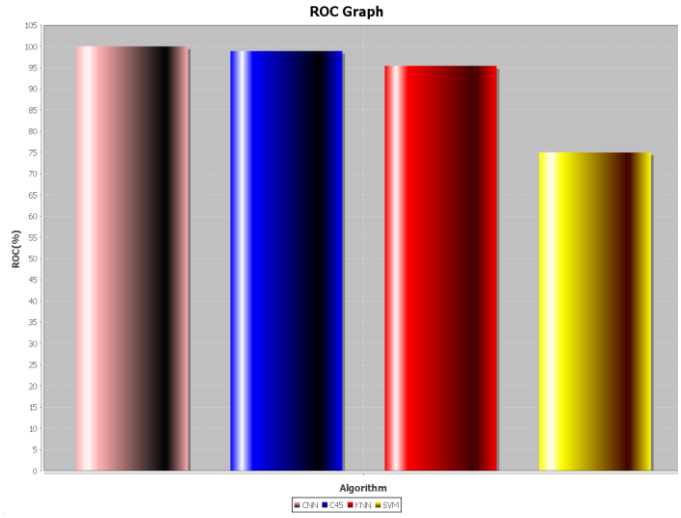

Figure .4. ROC Graph

V.CONCLUSION AND FUTURE SCOPE

Breast cancer is the most common disease and is usually associated with poor prognosis. Now currently needs to develop efficient computational techniques for the breast cancer prognosis prediction. The success stories of Deep Architectures, ANN (Artificial Neural Network) and Reinforcement Learning (RL) are used to making of the machines more intelligent. Here, a deep neural network is used to a prognosis of prediction the women breast cancer. A DNN model is composed of an input layer, multiple hidden layers, and output layers. Units between layers are fully connected.SVM, KNN and C 4.5 algorithms are used to the classification. For every round, nine of these ten subsets are further dividing into the training and validation set, while remaining one subset is for testing. Decision level multimodal fusion is used to integrate both clinical information and breast cancer-specific relationship between genes.

REFERENCES

[1] J. Xu et al., "Stacked sparse autoencoder (ssae) for nuclei detection of breast cancer histopathology images", IEEE Trans. Med. Imag., vol. 35, no. 1, pp.119-130,[2016].

[2] Y. Xu, Q. Feng, M. Lai, and E. I. C. Chang, "deep learning of feature representation with multiple instance learning for medical image analysis", in Proc. ICASSP, pp. 1626-1630, [2014].

[3] C. L. Chen et al., "deep learning in label-free cell classification", Sci. Rep., vol. 6, no. 1, Artical No. 2147, [2016].

[4] F. Ning, D. Delhomme, Y. LeCun, F. Piano, L. Bottou, and P. E. Barbano, "Toward automatic phenotyping of developing embryos from videos", IEEE Transactionon Image Processing, Volume 14, No. 9, pp. 1360-1371, [2005].

[5] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "mitosis detection in breast cancer histology images with deep neural networks", in Proc.MICCAI, 2013, pp. 411-418.

[6] M. Vivareli and A. D. Pinna,"ARTIFICIAL NEURAL NETWORK IS SUPERIOR TO MELD IN PREDICTING MORTALITY OF PATIENTS WITH END-STAGE LIVER DISEASE", Volume 56. No. 2. pp. 253–258. [2007].

[7] T. Parnamaa and L. Parts, "ACCURATE CLASSIFICATION OF PROTEIN SUBCELLULAR LOCALIZATION FROM HIGH-THROUGHPUT MICROSCOPY IMAGES USING DEEP LEARNING", G3, Volume 7, No. 5, pp. 1385-1392, [2017].

[8] A. Ferrari, S. Lombardi, and A. Signoroni, "BACTERIAL COLONY COUNTING WITH CONVOLUTIONAL NEURAL NETWORKS IN DIGITAL MICROBIOLOGY IMAGING", Pattern Recognition, Volume-61, pp. 629-640, [2017].

[9] Ao Li, Chen Peng and Ya Zhang, "Improve glioblastoma multi-forme prognosisprediction by using feature selection and multiple kernels learning", IEEE/ACMtransactions on computational biology and bioinformatics, 07 April 2016.