Convolutional Neural Network Based Image Classifier for Breast Cancer Histopathology Images

By

Md. Ashfakur Rahman Arju
ID: 173-25-631

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science and Engineering.

Supervised By

Md Zahid Hasan
Assistant Professor & Associate Head (In-Charge)
Department of CSE Daffodil
International University

DAFFODIL INTERNATIONAL UNIVERSITY

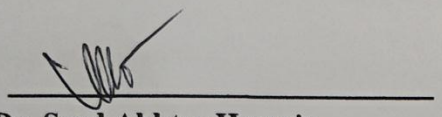
DHAKA, BANGLADESH

November, 2019

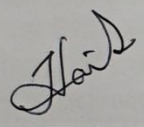
APPROVAL

This Thesis titled **"Convolutional Neural Network Based Image Classifier for Breast Cancer Histopathology Images"**, submitted by Md. Ashfakur Rahman Arju (ID:173-25-631) to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of MSc in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 28th November, 2018.

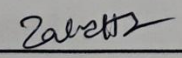
BOARD OF EXAMINERS



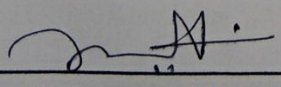
Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Dr. Sheak Rashed Haider Noori — **Internal Examiner**
Associate professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Md Zahid Hasan — **Internal Examiner**
Assistant Professor & Coordinator of MIS
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

I declare that, this project has been done by me under the supervision of **Md Zahid Hasan, Assistant Professor, Department of CSE,** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Md Zahid Hasan
Assistant Professor and Coordinator of MIS
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Submitted by:

Md. Ashfakur Rahman Arju
ID: 173-25-631
Department of CSE,
Daffodil International University

ACKNOWLEDGEMENT

First of all, my heartiest thanks and gratefulness to Almighty Allah for His divine blessing that makes us capable to complete this thesis successfully.

I would like to thank my honorable teacher & project supervisor **Md Zahid Hasan, Assistant Professor, Department of CSE,** Daffodil International University for his endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain**, Head, Department of CSE, for his kind help to finish our project and we are also thankful to all the other faculty and staff members of our department for their co-operation and help.

We must acknowledge with due respect the constant support and patients of our parents.

Finally, we would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

ABSTRACT

This project is on "**Convolutional Neural Network Based Image Classifier for Breast Cancer Histopathology Images"**. Thousands of people around the globe die every year as result of Breast Cancer (BC). Breast Cancer stages are ranging from early curable stages to late metastatic stages. An early detection of breast cancer can save millions of lives each year. Cancer cells can be detected through several ways like breast MRI scan, Mammogram, breast Ultrasound and Histopathology images. In our research we have used publicly available breast cancer histopathology image dataset hosted at http://web.inf.ufpr.br/vri/breast-cancer-database. We have tried to develop automated malignant and benign breast cancer detection system which can detect cancer from the histopathology images and thus making it more efficient and diagnosis more scalable and less prone to error. This type of research can be extended further to apply on other type of cancer detection. We have chosen Convolutional Neural Network (CNN) as our choice of classifier as recent research have shown that CNN has the most upper hand in compare to others hand crafted feature descriptors when classifying medical images. But for the CNN the computation time and sophistication required is very high.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Cancer is a significant public health problem in the world today. According to the IARC (International Agency for Research on Cancer) of the WHO (World Health Organization), 8.2 million deaths were caused by cancer in 2012 and 27 million of new cases of this disease are expected before 2030 [1]. In particular, breast cancer (BC) is one of most common type of cancer among women.

Mortality of BC is very high when compared to other types of cancer. Detection and diagnosis of BC can be achieved by imaging procedures such as diagnostic mammograms (x-rays), magnetic resonance imaging, ultrasound (sonography), and thermography [2]. Imaging for cancer screening has been investigated for more than four decades [3]. However, biopsy is the only way to diagnose with confidence if cancer is really present. Among biopsy techniques, the most common are fine needle aspiration, core needle biopsy, vacuum-assisted and surgical (open) biopsy (SOB) [4]. The procedure consists in collecting samples of cells or tissue, which are fixed across a glass microscope slide for subsequent staining and microscopic examination. Diagnosis from a histopathology image is thus the gold standard in diagnosing almost all types of cancer, including BC [5], [6]. The final BC diagnosis, including grading and staging, is done by pathologists applying visual inspection of histological samples under microscope.

1.1 Objectives

The main objective of this thesis is to develop an automated agent which can successfully predict benign and malignant cancer from different type of images from BreakHis [10] dataset.

The goals of our thesis are:

- To reshape differently sized images and convert them into NPZ arrays.
- Successfully divide the dataset into test and train folders.
- Train AlexNet with existing model weights.
- Retrain AlexNet with our new train dataset.
- Save model progress as model checkpoints
- Save best model weights for future prediction
- Predict breast cancer type from unknown test dataset.

- Plot cancer type based on best probabilistic match from the training weights.

1.2 Motivation

Breast Cancer (BC) is one of the most prevailing cause of death among women. Every year millions of women die around the globe from different type of cancer. Detection and classification of cancer type is a very sophisticated and time-consuming process which requires expert pathologist with years of experience. Moreover, human being is prone to fatigue and error.

Our motivation was to develop and train an agent which can classify cancer cells from histopathology images. The learned agent will reduce the sophistication and time required to classify breast cancer type thus helping patients with higher accuracy. Moreover, an automated agent can predict and classify images at a greater rate and higher accuracy reducing the amount of man power.

1.3 Expected Outcome

In our Thesis our main focus is to develop and train an agent which can predict cancer type from 7900 images with different magnification factors. The agent is based on convolutional neural network where we have used a pre-trained model (AlexNet) rather than a hand crafted one. AlexNet has a performance boost of 10% in comparison with hand crafted models. After the training the agent has the accuracy of 97% and can predict benign and malignant breast cancer from 40x, 100x, 200x, 400x magnified images.

1.4 Report Layout

First Chapter contains the Introduction, Objectives, Motivation, Expected Outcome and Report layout of our project. Then second chapter contains Project Introduction, Related works, Comparative Studies, Scope of the problem and also Challenges of our project. Third chapter contain all about Requirement Specification which are Use Case Modeling and Description, Logical Data Model, Design Requirements and Description of the Dataset we have used.

Fourth Chapter describes Proposed methodology in details and Training strategies for the agent. Our fifth chapter is all about Implementation and Accuracy testing. This contains Implementation of AlexNet, Prediction Visualizer and Testing modules.

Our last chapter contain conclusion of the full thesis. This report contains all about our proposed system, its problem, solution and future improvements.

CHAPTER 2

BACKGROUND AND RELATED WORKS

2.1 Introduction

Thousands of people around the globe die every year as result of Breast Cancer (BC). Breast Cancer stages are ranging from early curable stages to late metastatic stages. An early detection of breast cancer can save millions of lives each year. Cancer cells can be detected through several ways like breast MRI scan, Mammogram, breast Ultrasound and Histopathology images. In our research we have used publicly available breast cancer histopathology image dataset hosted at http://web.inf.ufpr.br/vri/breast-cancer-database. We have tried to develop automated malignant and benign breast cancer detection system which can detect cancer from the histopathology images and thus making it more efficient and diagnosis more scalable and less prone to error. This type of research can be extended further to apply on other type of cancer detection. We have chosen Convolutional Neural Network (CNN) as our choice of classifier as recent research have shown that CNN has the most upper hand in compare to others hand crafted feature descriptors when classifying medical images. But for the CNN the computation time and sophistication required is very high. Mortality of BC is very high when compared to other types of cancer. Detection and diagnosis of BC can be achieved by imaging procedures such as diagnostic mammograms (x-rays), magnetic resonance imaging, ultrasound (sonography), and thermography. Imaging for cancer screening has been investigated for more than four decades. However, biopsy is the only way to diagnose with confidence if cancer is really present. Among biopsy techniques, the most common are fine needle aspiration, core needle biopsy, vacuum-assisted and surgical (open) biopsy (SOB). The procedure consists in collecting samples of cells or tissue, which are fixed across a glass microscope slide for subsequent staining and microscopic examination. Diagnosis from a histopathology image is thus the gold standard in diagnosing almost all types of cancer, including BC. The final BC diagnosis, including grading and staging, is done by pathologists applying visual inspection of histological samples under microscope.

2.2 Automated Agent Scenario

Millions of people around the world are diagnosed with Breast Cancer (BC) each year. The process of diagnosing BC needs very sophisticated machineries and experienced pathologist. Moreover, human is prone to fatigue and has a limitation of how much work the can do each day. To overcome his type of limitations an automated agent can be of great help. Machines are free from fatigue and have no limitation of working hours. A trained agent can predict quicker than human and can process large amount of data. New classifications can be programmed to be trained again which in terms increases agent's accuracy.

2.3 Saved Model and Reuse

Models can be saved after the training is done which can be reused for latter model training. Reusing model weights reduces huge time required to train an agent. we reuse a recently built model design and the vast majority of the learned weights, and after that utilization standard preparing strategies to take in the remaining, non-reused parameters. When you build your Keras display utilizing the useful interface, you can likewise assemble extra models on any subset of the ways through the system by reusing the go-between capacities. At that point you can prepare on just parts of the system (given that you have focuses for the yields). I haven't endeavored to prepare on sub-systems of a system, however I do utilize these middle person models to engender enactments between interior layers. The Python bundle conx that is based over Keras will construct these middle person models for you, coincidentally.

2.4 Related Works

The automatic imaging process for cancer diagnosing has been explored as a subject of analysis for over forty years [3] however remains difficult thanks to the quality of the images to analyze. For example, Kowal et al. [11] compare and test different algorithms for nuclei segmentation, where the cases are classified as either benign or malignant on a dataset of 500 images, and report accuracies ranging from 96% to 100%. Filipczuk et al. [12] present a BC diagnosis system based on the analysis of cytological images of fine needle biopsies, to discriminate the images as either benign

or malignant. Using four completely different classifiers trained with a 25-dimensional feature vector, they report a performance of ninety-eight on 737 pictures.
Until recently, most of the works on BC histopathology image analysis were carried out on small datasets, which are usually not available to the scientific community. Contributing to mitigate this gap, Spanhol et al. [20] introduced a dataset composed of 7,909 breast histopathological images acquired on 82 patients. In the same study, the authors evaluated six completely different textural descriptors and different classifiers and reported a series of experiments with accuracy rates starting from eightieth to eighty fifth, depending on the image magnification factor. Based on the results presented in [21], it is undeniable that the texture descriptors can offer a good representation to train classifiers. However, some researchers advocate that the main weakness of the current machine learning methods lies exactly on this feature engineering step [22], [23]. To them, machine learning algorithms should be less dependent on feature engineering by being able to extract and organize the discriminative information from the data, in other words, should be capable of learning the representation.
A considerable amount of efforts has thus been devoted to the field of BC histopathology image analysis, and in particular to the automated classification of benign or malignant images, for computer-aided diagnosis. Kowal et al. [9] compare and test different algorithms for nuclei segmentation on a dataset of 500 images, for which accuracies ranging from 96% to 100% are reported. Filipczuk et al. [10] present a BC diagnosis system based on the analysis of cytological images of fine needle biopsies, to discriminate the images as either benign or malignant. Using four different classifiers trained with a 25-D feature vector, they report a performance of 98% on 737 images. Similarly, to [9] and [10], George et al. [11] propose a diagnosis system for BC
based on the nuclei segmentation of cytological images. Using different machine learning models, such as neural networks and support vector machines (SVMs), they report accuracy rates ranging from 76% to 94% on a dataset of 92 images. Zhang
et al. [12] propose a cascade approach with rejection option. In the first level of the cascade, authors expect to solve the easy cases, while the hard ones are sent to a second level where a more complex pattern classification system is used.
Most of these recent works related to BC classification are focused on Whole-Slide Imaging (WSI) [18], [17], [16]. However, the broad adoption of WSI and other forms

of digital pathology still facing obstacles such as the high cost of implementing and operating the technology, insufficient productivity for high-volume clinical routines, intrinsic technology-related considerations, unsolved regulatory issues, as well as "cultural resistance" from the pathologists [19].

Similarly, to [13] and [14], George et al. [15] propose a diagnosis system for BC based on the nuclei segmentation of cytological images. Using totally different machine learning models, such as neural networks and support vector machines, they report accuracy rates ranging from 76% to 94% on a dataset of 92 images. Zhang et al. [16] propose a cascade approach with rejection option. In the first level of the cascade, authors expect to solve the easy cases while the hard ones are sent to a second level where a more complex pattern classification system is used. They assess the projected technique on a information projected by the Israel Institute of Technology, that consists of 361 pictures and report results of ninety seven of responsibility. In another work [17], the same authors assess an ensemble of one-class- classifiers on the same database achieving a recognition rate of 92%. They assess the proposed method on a database proposed by the Israel Institute of Technology, which is composed of 361 images (40×magnification). On this dataset, they report results of 97% of reliability. In another work [13], the same authors assessed an ensemble of one-class classifiers on the same database achieving a recognition rate of 92%. We can gather from the literature that most of the works on BC histopathology image analysis are carried out on small datasets, which are usually not available to the scientific community. In a recent review, Veta et al. [14] point out that the main obstacle in the development of new histopathology image analysis methods is the lack of large, publicly available, annotated datasets. Annotated database is also crucial to develop and validate machine learning systems.

The idea of representation learning is not new but it emerged only recently as a viable alternative due to the appearance and popularization of the Graphic Processing Units (GPUs) which are capable of delivering high computational throughput at relatively low cost, achieved through their massively parallel architecture. Among the different approaches, the Convolutional Neural Network (CNN) introduced by LeCun in [24], has been widely used to achieve state-of-the-art results in different pattern recognition problems [25], [26]. In the case of texture classification it has not been different. Hafemann et al. [27] have shown, for images of microscopic and macroscopic texture, that CNN is able to surpass traditional textural descriptors. Besides, the traditional

approach to extract appropriate features for classification tasks in pathological images requires considerable efforts and effective expert domain knowledge, frequently leading to highly customized solutions, specific for each problem and hardly applicable in other contexts [28].

Our proposed work is based on two previous work Fabio A. Spanhol and Luiz S. Oliveira Method and Caroline Petitjean and Laurent Heutte Method. A comparative study of the two method is described below.

2.4.1 Fabio A. Spanhol and Luiz S. Oliveira Method

Fabio A. Spanhol and Luiz S. Oliveira has proposed 1-NN (Nearest Neighbor) classifier-based approach. They have used several feature extractors such as Local Binary Patterns (LBP), Completed Local Binary Pattern (CLBP), Local Phase Quantization (LPQ), Gray-Level Co-Occurrence Matrices (GLCM), Parameter-Free Threshold Adjacency Statistics (PFTAS), Oriented FAST and Rotated BRIEF (ORB) etc. Their methodology has an overall accuracy of 91.8%. Accuracy distribution of the methodology can be visualized from the following figure 1

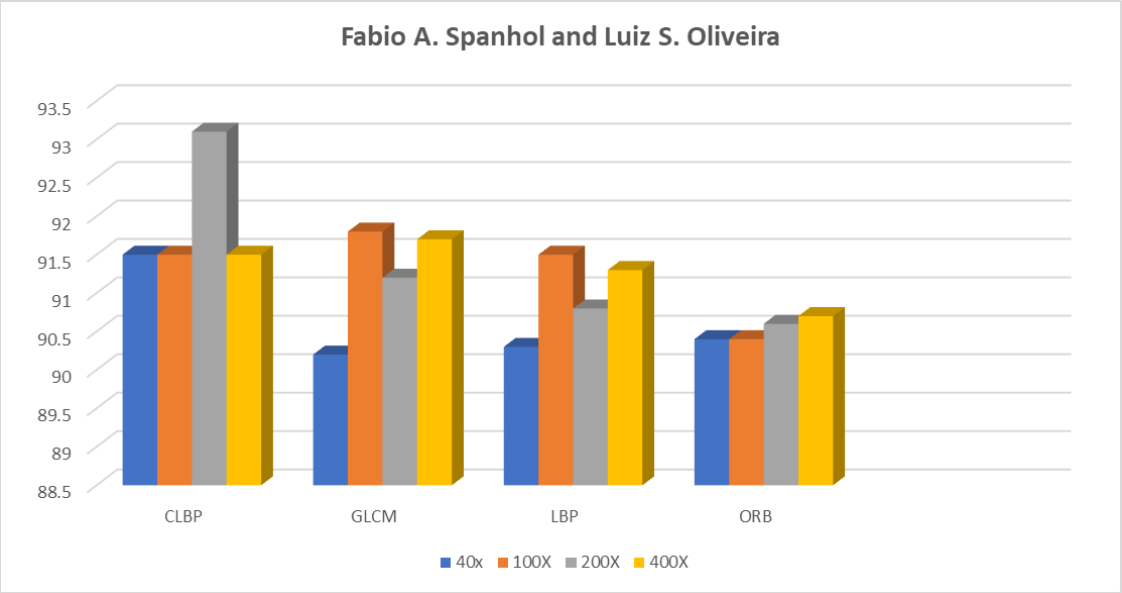


Figure 2.1: Fabio A. Spanhol and Luiz S. Oliveira Method Accuracy

2.4.2 Caroline Petitjean and Laurent Heutte Method

Caroline Petitjean and Laurent Heutte has proposed Random Forests (RF) classifier-based approach. They have used several feature extractors such as Local Binary Patterns (LBP), Completed Local Binary Pattern (CLBP), Local Phase Quantization (LPQ), Gray-Level Co-Occurrence Matrices (GLCM), Parameter-Free Threshold Adjacency Statistics (PFTAS), Oriented FAST and Rotated BRIEF (ORB) etc. Their methodology has an overall accuracy of 89.9%. Accuracy distribution of the methodology can be visualized from the following figure 2

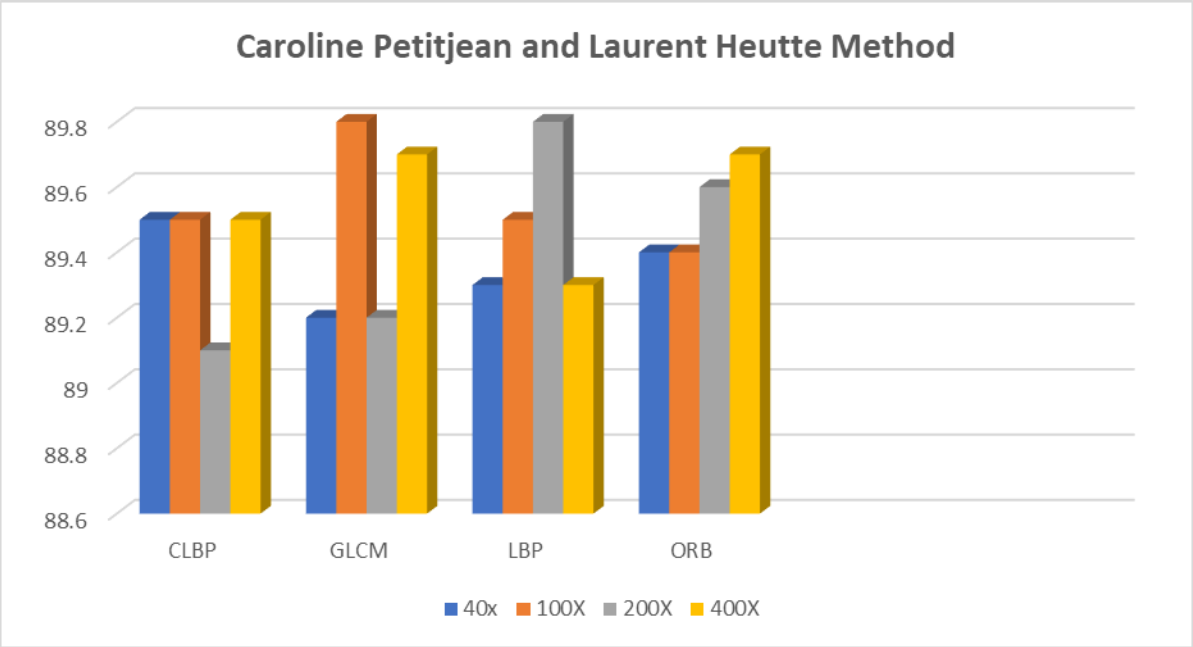


Figure 2.2: Caroline Petitjean and Laurent Heutte Method Accuracy

2.5 Comparative Studies

Fabio A. Spanhol and Luiz S. Oliveira work is based on SVM classifier (Support Vector Machine) and Caroline Petitjean and Laurent Heutte work is based on Random Forest (RF). Both of the above methods have used same feature extractors like Local Binary Patterns (LBP), Completed Local Binary Pattern (CLBP), Local Phase Quantization (LPQ), Gray-Level Co-Occurrence Matrices (GLCM), Parameter-Free Threshold Adjacency Statistics (PFTAS), Oriented FAST and Rotated BRIEF (ORB) etc.

Random Forest is best for multiclass problems, while SVM can only predict two-class. For multiclass problem dataset needs to be reduced into multiple binary classification problems.

Random Forest works well with dataset having mixed features. Random Forest treats data as it is. SVM relies on the distance between different points. It is difficult to decide if distance is meaningful or not. Because of this constrain, features need to be one-hot encoded before processing. Moreover, different scaling like min, max, average etc. are highly recommended before preprocessing step.

2.6 Scope of the Problems

Main scope of this thesis is as follows

1. Develop and train an agent which can predict Breast Cancer (BC) from more than 8000 images based on BreakHis dataset.
2. Save best weights from each iteration for future use which can in terms reduce time required for prediction.

2.7 Challenges

Throughout the work we have faced several challenges. Most prominent challenges are stated below

1. **Overfitting**

   The main challenges of this thesis were to reduce overfitting throughout the training epochs. Sometimes the agent performed very poor on the unseen data. To reduce this type of problem we have followed the several steps like cross-validation, increasing training data volume, reduce features, regularization, ensembling and early stopping.

2. **Selection of Activation Function**

   We have tried several activation functions like Binary Step, Sigmoid, Tanh, ReLU, Leaky ReLU, Softmax etc for the pullout layer of our CNN model. Though each activation function has its strong points, ReLu works best for the AlexNet architecture and have the best weight distribution.

3. **Distribution of Tensors**

We have used Nvidia Quadro K200m GPU for our work, which in term is a very low power GPU of only having 1.67GB of video memory. This low memory was very frustrating while assigning tensors as it always ran out memory all time. To overcome this problem, we have used fusion of both GPU and CPU while assigning tensor. We were needed to active multiple CPU workers at a time.

4. **Choosing the Best Weights**

In each epoch model generates several weights it has trained on. Choosing the best weights in the past was bit of challenging. In modern days several frameworks are use to choose best weights from the several iterations in each epoch. Keras worked best in this type of scenario to the best of knowledge.

CHAPTER 3

REQUIREMENT SPECIFICATION

3.1 Introduction to Dataset

The BreaKHis database [29] contains microscopic biopsy images of benign and malignant breast tumors. Images were collected through a clinical study from January 2014 to Dec 2014. All patients stated the P&D research laboratory, Brazil, during this period of time, with a clinical indication of BC were invited to participate in the study. The institutional review board approved the study and all patients gave written informed consent. All the data were anonymized.

Samples are generated from breast tissue diagnostic test slides, stained with hematoxylin and eosin (HE). The samples are collected by surgical (open) biopsy (SOB), prepared for histological study and labeled by pathologists of the P&D Lab. The preparation procedure employed in this work is that the customary paraffin method, which is widely used in clinical routine. The main goal is to preserve the initial tissue structure and molecular composition, allowing to observe it in a light microscope. The complete preparation procedure includes steps such as fixation, dehydration, clearing, infiltration, embedding, and trimming [30]. To be mounted on slides, sections of around 3µm are cut using a microtome. After staining, the sections are covered with a glass coverslip. Then the pathologists identify the tumoral areas in each slide, by visual analysis of tissue sections under a microscope. Final identification of every case is made by skilled pathologists and confirmed by complementary exams like assay (IHC) analysis.

An Olympus BX-50 system microscope with a relay lens with magnification of 3.3× coupled to a Samsung digital color camera SCC-131AN is used to obtain digitized images from the breast tissue slides. Images are nonheritable in 3-channel RGB (Red-Green-Blue) TrueColor (24-bit color depth, eight bits per color channel) color area exploitation magnifying factors of 40×, 100×, 200× and 400×, similar to objective lens 4×,10×, 20×, and 40×.

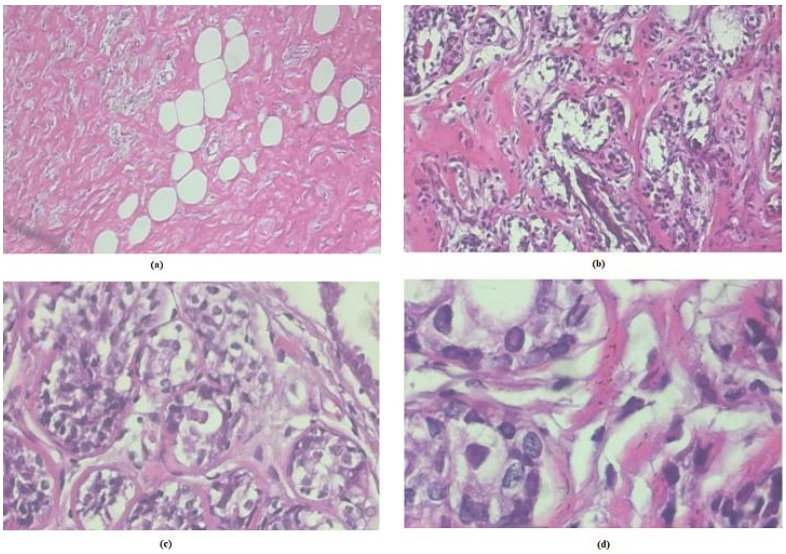


Figure 3.1: A slide of breast malignant tumor (stained with HE) seen in different magnification factors: (a) 40×, (b) 100×, (c) 200×, and (d) 400×.

Table I

IMAGE DISTRIBUTION BY MAGNIFICATION FACTOR AND CLASS

Magnification	Benign	Malignant	Total
40×	625	1,370	1,995
100×	644	1,437	2,081
200×	623	1,390	2,013
400×	588	1,232	1,820
Total	2,480	5,429	7,909
# Patients	24	58	82

3.2 Workflow of the Proposed Method

We have used Convolutional Neural Network (CNN) based on AlexNet architecture. CNN has several modules like input layer, convolutional layer, pooling layer, ReLU and fully connected layer. We have used AlexNet as CNN model as it has a fair prediction accuracy over CNN models like VVG Net, ResNet 10, ResNet 50 etc.

Our dataset contains more than 7500 images of benign and malignant breast cancer. Each cancer type Histopathological image has zoom factor of 40x, 100x, 200x, 400x and are different resolution. Before training CNN model, we must reshape all the images to the same resolution.

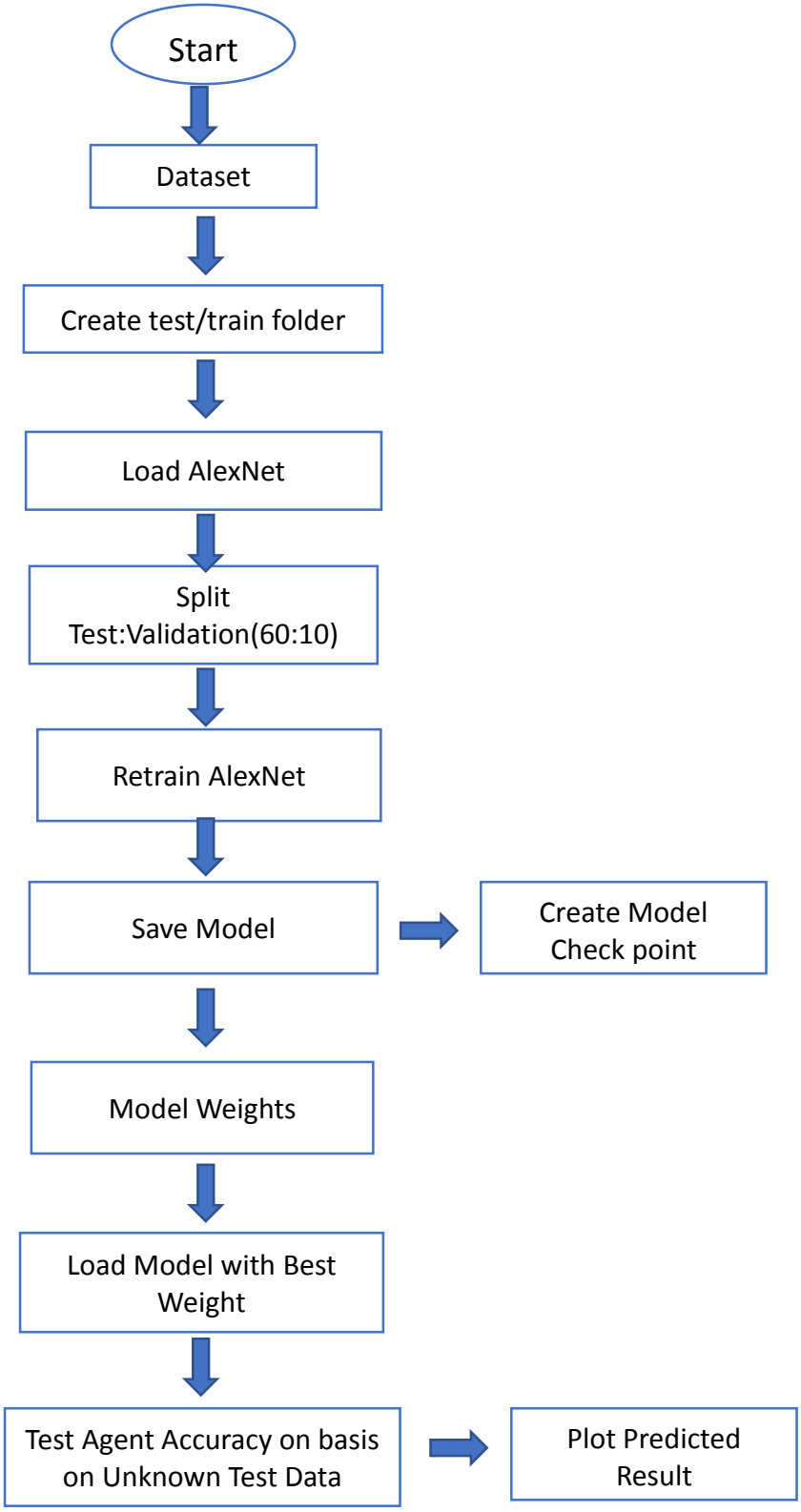


Figure 3.2: Proposed methods work flow

After reshaping the images, we have converted them into SciKit arrays known as NPZ arrays to reduce huge time required to train high resolution images. This step adds minor execution time but reduces overall system execution time. We feed the NPZ arrays to the input layer but before that we need load empty AlexNet model. From the dataset we have created test and train folder and placed random images. Train folder is used to train the model and test folder serve as the unknown images source. We have trained the AlexNet with the train dataset which contains 60% of total images and each training iteration is validated against 10% of total images known as validation set. After retraining the AlexNet with current train dataset we save the best weights from each iteration and save model state. Saved state can be later used to retrain where left off. We have chosen the best weight from almost 1000 weights to predict the unseen data from the test dataset. Test dataset contains 30 % of total images which are unseen by the model. After the prediction we plot the best probabilistic prediction on the unseen image. The whole work flow can be visualized from the figure 4.

3.3 Workflow Graph of the Proposed Method

BreakHis dataset contains more than 7500 images of different size. Dataset images are categorized as benign and malignant type with subclass of adenosis, ductal carcinoma, fibroadenoma, lobular carcinoma, mucinous carcinoma, papillary carcinoma, phyllodes tumor, tubular adenoma. Each subclass has zoom factor of 40x, 100x, 200x, 400x. AlexNet only accept 256*256 images so we need to reshape the images of the train data set. Reshaping and scaling of the images are done in the input layer. RGB color channel allocation are done in this layer to. In our case we have taken 3 channels. In feature learning process there are three parts. Convolution layer extracts the high-level features of each images from the input images. After we have extracted high level features from the input images, we apply ReLu (Non-Linear Rectified Unit) on each convolution layer immediately the purpose of this layer is to introduce nonlinearity to a system that basically has just been computing linear operations during the conv layers (just element wise multiplications and summations).

After the ReLu we apply max poling. Max pooling layer chooses the best features from the primary features extracted by the convolution layer. Max pooling gives us the best features which are multidimensional arrays. As our fully connected layers only learns on single dimensional array, we need to platen the multidimensional array before feeding to the fully connected layers. Fully connected layers learn on the flatten inputs by applying the back propagation. For the back propagation and distribution of the images we have used ADAM function rather than Stochastic Gradient Descent (SGD). ADAM is much more optimized and reduces compilation time. Fully connected layers outputs an N dimensional vector where N is the number of classes that the program has to choose from. Each number of this N dimensional vector represents the probability of a certain class.

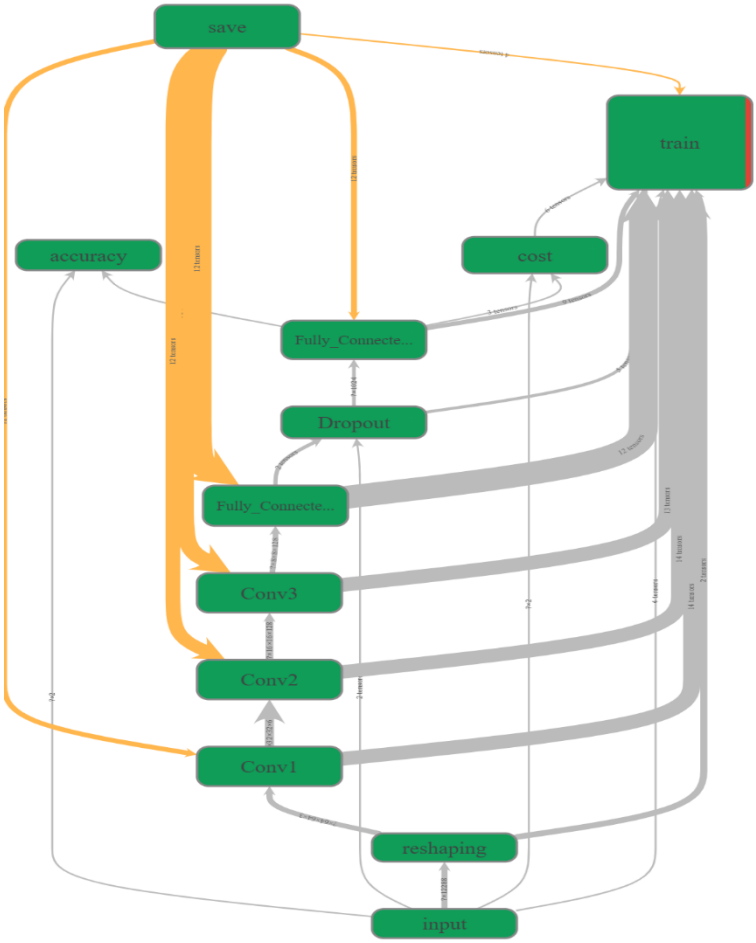


Figure 3.3: System Workflow Graph

We have ventured classification problem rather than localization and detection through our thesis last layer of the model is the Softmax function. Softmax assigns

decimal probability to each class in the multi-dimensional feature array. Total probability of the Softmax function assigned to each class must be 1. Softmax also helps the model training to converge more quickly which will take much longer without the Softmax.

3.4 Prediction flow of the Proposed Method

We have generated the best weight of each iteration. After the model is finished learning with all the images in the train dataset the model will generate more than 100 weights. We use this weight to predict the class the unknown image belongs to. First the model is loaded with the best weight. We then fed this model a new unknown image from the test dataset. Trained model predicts the class of the unknown image. After the probabilistic prediction we plot the result with the help of image plotter. CNN models compare each prediction with the ground truth. Image plotter prints the highest probability and the nest best match on the unseen image.

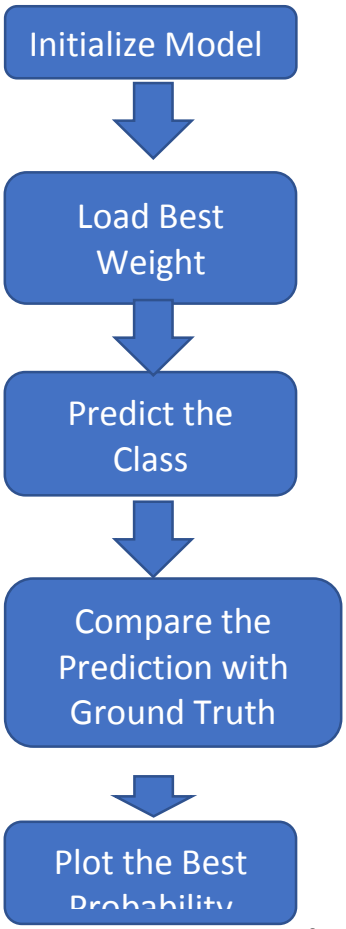


Figure 3.4: Prediction Workflow

3.5 Implementation Requirements

We have used several frameworks and python as programming language. Frameworks we have used through the thesis are listed below

- **Keras**

  Keras is AN open supply neural network library written in Python.

  It is capable of running on prime of TensorFlow, Microsoft psychological feature Toolkit, or Theano. Designed to alter quick experimentation with deep neural networks, it focuses on being easy, modular, and extensile. It was developed as part of the research effort of project ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System) and its primary author and maintainer is François Chollet, a Google engineer.

  In 2017, Google's TensorFlow team set to support Keras in TensorFlow's core library.Chollet explained that Keras was planned to be AN interface instead of a standalone machine-learning framework. It offers a higher-level, additional intuitive set of abstractions that create it straightforward to develop deep learning models despite the process backend used. Microsoft superimposed a CNTK backend to Keras similarly, offered as of CNTK v2.0.

- **Tensorflow**

  TensorFlow is associate open supply code library for numerical computation mistreatment information flow graphs. The graph nodes represent mathematical operations, whereas the graph edges represent the multidimensional information arrays (tensors) that flow between them. this versatile design allows you to deploy computation to at least one or additional CPUs or GPUs in a very desktop, server, or mobile device while not revising code. TensorFlow additionally includes TensorBoard, a knowledge mental image toolkit.

  TensorFlow was originally developed by analyzers and engineers performing on the Google Brain team at intervals Google's Machine Intelligence analysis organization for the needs of conducting machine learning and deep neural networks research. The system is general enough to be applicable in a very big variety of different domains, as well.

TensorFlow provides stable Python API and C Apis furthermore as while not API backwards compatibility guarantee like C++, Go, Java, JavaScript and Swift.

- **SciKit Learn**

  Scikit-learn (formerly scikits.learn) may be a free computer code machine learning library for the Python programing language. It options numerous classifications, regression and bunch algorithms together with support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is intended to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

  The scikit-learn project started as scikits.learn, a Google Summer of Code project by David Cournapeau. Its name stems from the notion that it's a "SciKit" (SciPy Toolkit), a separately-developed and distributed third-party extension to SciPy. the initial codebase was later rewritten by different developers. In 2010 Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort and Vincent Michel, all from INRIA took leadership of the project and created the primary public unharness on Feb the first 2010.Of the assorted scikits, scikit-learn likewise as scikit-image were delineate as "well-maintained and popular" in November 2012. As of 2018, scikit-learn is in active development.

- **OpenCV**

  OpenCV (Open supply pc vision) may be a library of programming functions in the main geared toward period pc vision. Originally developed by Intel,it was later supported by Willow Garage then Itseez (which was later nonheritable by Intel). The library is cross-platform and free to be used beneath the ASCII text file BSD license.OpenCV supports the deep learning frameworks TensorFlow, Torch/PyTorch and Caffe.

  Officially launched in 1999, the OpenCV project was ab initio associate Intel analysis initiative to advance CPU-intensive applications, a part of a series of comes together with period ray tracing and 3D show walls. The main contributors to the project enclosed variety of optimization consultants in Intel Russia, also as Intel's Performance Library Team. within the period of OpenCV, the goals of the project were described as:

Advance vision analysis by providing not solely open however conjointly optimized code for basic vision infrastructure. No additional reinventing the wheel. Disseminate vision data by providing a standard infrastructure that developers may ride, so code would be additional promptly clear and transferable. Advance vision-based industrial applications by creating moveable, performance-optimized code accessible without charge – with a license that failed to need code to be open or free itself.

CHAPTER 4

PROPOSED WORK

Deep learning explores the chances of learning features directly from input images, avoiding hand-crafted models. The key concept of deep learning is to explore multiple levels of illustration aiming that higher-level features represent an abstract view of the images. Convolutional Neural Networks (CNNs) now a days is used in everywhere from medical images classification to object localization. CNN is constructed of multiple convolutional layers stacked on top of each other, followed by a supervised deep net known as fully connected layer and sets feature maps represent both input and output of each convolutional layers. Input may very like image, audio, and video. In our case we have used color images, at the input layer each feature map is a two-dimensional array storing RGB channel of the input image. Output from each layer consists of a set of arrays where feature map represents a particular feature extracted at a particular input layer. A deep net is trained by feeding it input and letting it compute layer-by-layer to generate the final output for comparison with the correct answer. ADAM function work as weight distributor in each iteration and error are back propagated to net. At each step backward, the model parameters are tuned in a direction that tries to reduce the error. This process increases model accuracy as the learning progresses. Generally, training is done by feeding the model train data set again and again in an iterative fashion until the model converges.

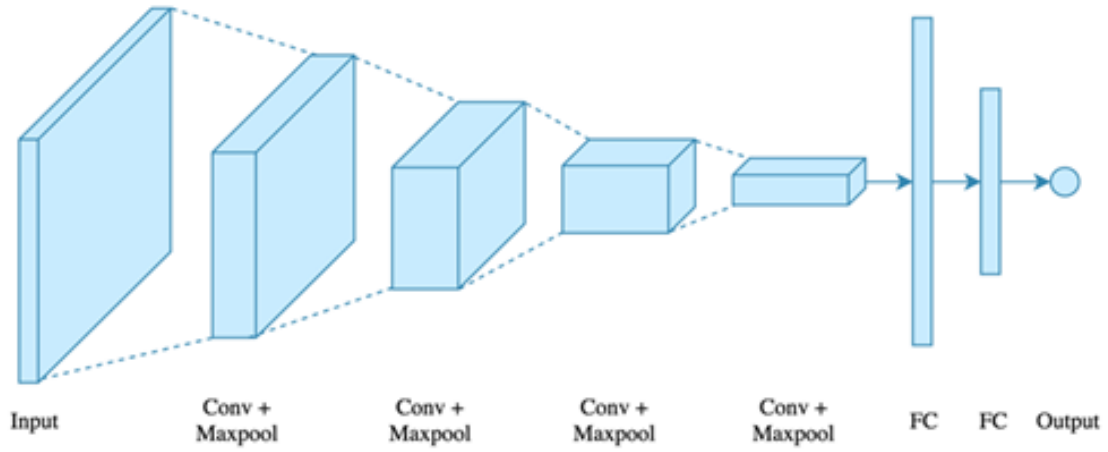


Figure 4.1: A CNN architecture

4.1 AlexNet Architecture

We have experimented several existing CNN models to classify breast cancer type from the BreakHis dataset. First, we started with VVG NET [31] a well-tested CNN model known to be working satisfactory for classifying digits from MINST [33] dataset. On MINST dataset VVG NET performs very well on classifying but performs below average than the already exiting works on classifying breast cancer type from the BreakHis data set. The best accuracy we have achieved with VVG NET is 73% at max where already existing works have achieved almost 90% accuracy.

While testing different existing models we have seen that only complex model like Inception V3, GoogleNet, AlexNet etc performs on best on classifying high resolution images. After bench marking the above-mentioned models, we have found that different variant of AlexNet [26] has sharp edge over other models in execution and performance. AlexNet was originally suggested by Alex Krizhevsky to classify images from the CIFAR-101 dataset which contains more than one million images of different size and more than 100 mutually exclusive classes. In CIFAR each class contains more than 1000 images. Architecture of AlexNet has the typical model architecture of a CNN model. The main differences are the deep connected layers and Softmax function at the end of the model. Generally, a CNN model consists of convolution layer, ReLu layer, max pooling layer, fully connected layer and probabilistic Softmax function as shown in the figure 6. AlexNet also has the same architecture but it has 3 convolution layers stacked one over other, 3 fully connected deep layer and probabilistic Softmax function for predicting the model outcome. AlexNet can perform very well on images that have wide variety of textures. Moreover, AlexNet support GPU acceleration which significantly increases the computational power reducing the required time to train and agent. A typical AlexNet have the following components

• Input layer: Input images are feed to this layers and output is feed to convolution layers. Reshaping and feature-scaling is done in this stage to prevent model from dimensionality error. BrekHis dataset contains images of different shape as result to correct dimensionality error reshaping is must. We have reshaped all the images to (64*64) pixel. As all the images are colored, we have chosen channel size to be 3 (RGB channel).

• Convolutional layers: Convolution layers convolves the input images into a multi-dimensional feature map with as set of learnable filters. We have used three convolutional layers. The kernels or filters are of size 5 × 5, padding is set to 2 and stride is set to SAME means stride and padding is of same size. The first two convolutional layers learn 32 filters each one and Gaussian distribution with standard deviation of 0.0001 and 0.01, respectively is initialized for convolutional layer 1 and 2 respectively. The last layer learns 64 filters and a Gaussian distribution with standard deviation of 0.0001 is used to initialize this layer.

• Pooling layers: Pooling layers are responsible for down- sampling the spatial dimension of the input. After each convolutional layer there is a polling layer. We have used a 2*2 kernel for each poling layer and a stride of size 2. The first pooling layer uses the max pooling over the generated feature map and the last two perform average pooling.

• ReLU layers: We have tested different activation functions like Tanh, Sigmoid ReLu, Leaky ReLu, Binary step etc. From our findings we have seen that ReLu works best onh the histopathology images. After each polling there is a ReLu layer. For an input value of x ReLu computes the neuron's output f (x) as x if x > 0 and (α × x) if x <= 0. α specifies whether to omit the negative part by multiplying it with the slope value (0.01 or so) rather than setting it to 0. The default value of α is 0. If the value of α is not set then it works as a standard ReLU function f (x) = max(0, x), in other words, the activation is simply threshold at zero.

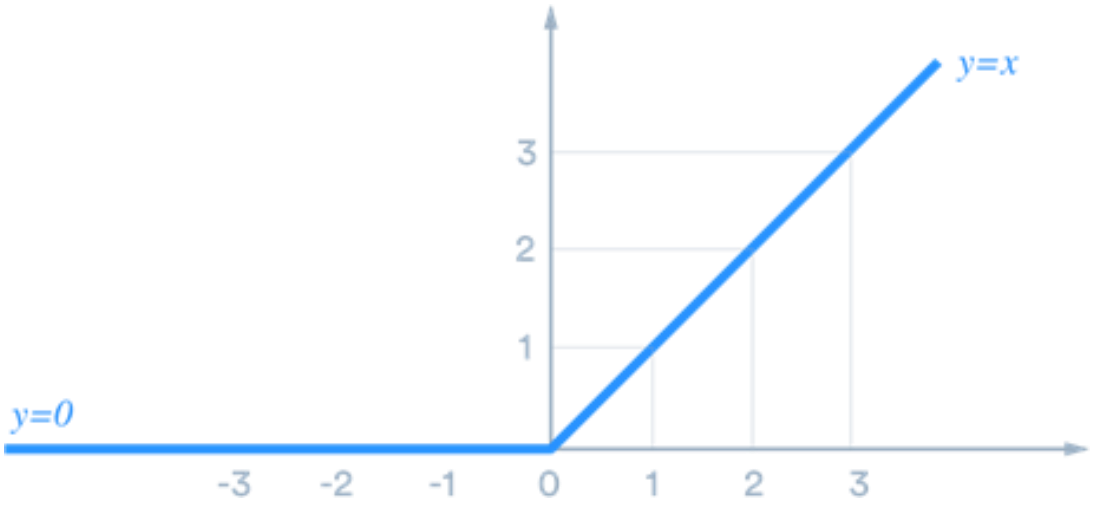


Figure 4.2: ReLU

• Fully connected layers: Fully connected layers are in theory a connected deep neural network that takes a multi-dimensional feature map as input and produces a single dimensional feature map as output. We have used three fully connected layers as this gives the best accuracy and optimum weight distribution. After the fully connected layer there is a probabilistic Softmax activation function. Softmax is a probabilistic function that gives the best match depending on the number of classes in the classification problem. For our case it's either benign or malignant breast cancer.

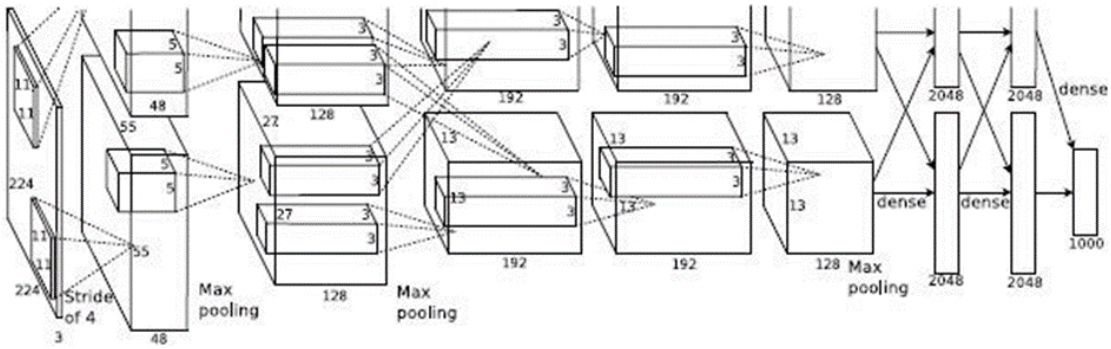


Figure 4.3. AlexNet CNN architecture.

4.2 Classification

To classify an image, we have combined the patch result for the whole image. We have divided the input images into because our model was trained on patches of images. So newly created patches were run through the model and the results were combined for the classification. We have extracted grid patches which has a good balance between classification problem and computational problem. Run the model with the best weights and the patches outputs the best probable match of the classes. We have used the Sum rule to combine the patch results which gives the best result in comparison to different fusion rules.

4.2 Developed Work

Following figures illustrates the different aspects of our developed work

- Converting Dataset into NPZ arrays

We have used NumPy to convert images into NPZ arrays. By converting the images into pixelated arrays reduces computation time significantly. Dataset directory is pasted into the CMD and underling python script converts all the images in to NPZ arrays.

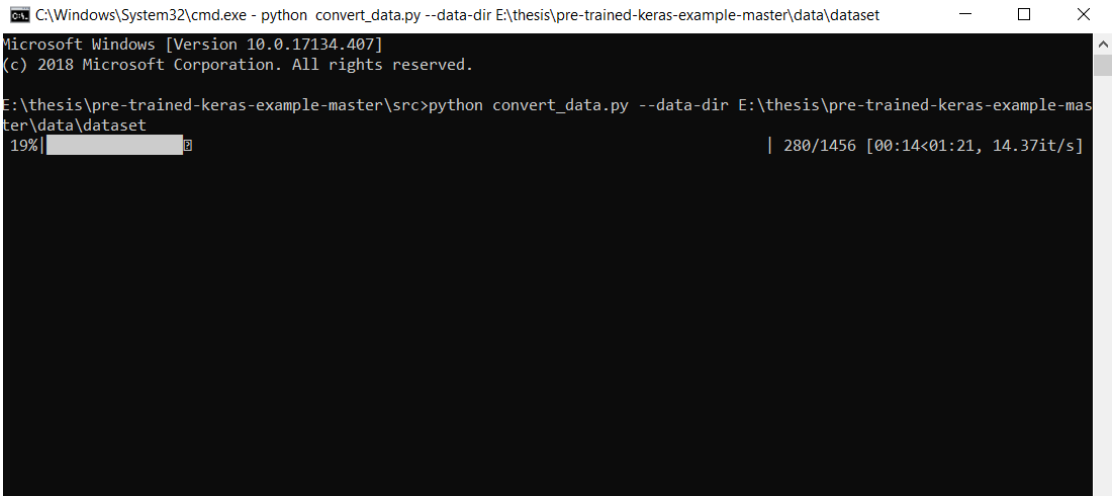


Figure 4.4: Converting Images into NPZ Arrays

- **Train AlexNet**

  After converting the dataset into NPZ arrays we feed to the AlexNet. AlexNet trains on the test dataset and saves the best weights in the weight directory for each iteration. While the model trains on large volume of data we need to save the model state to avoid any interruption.

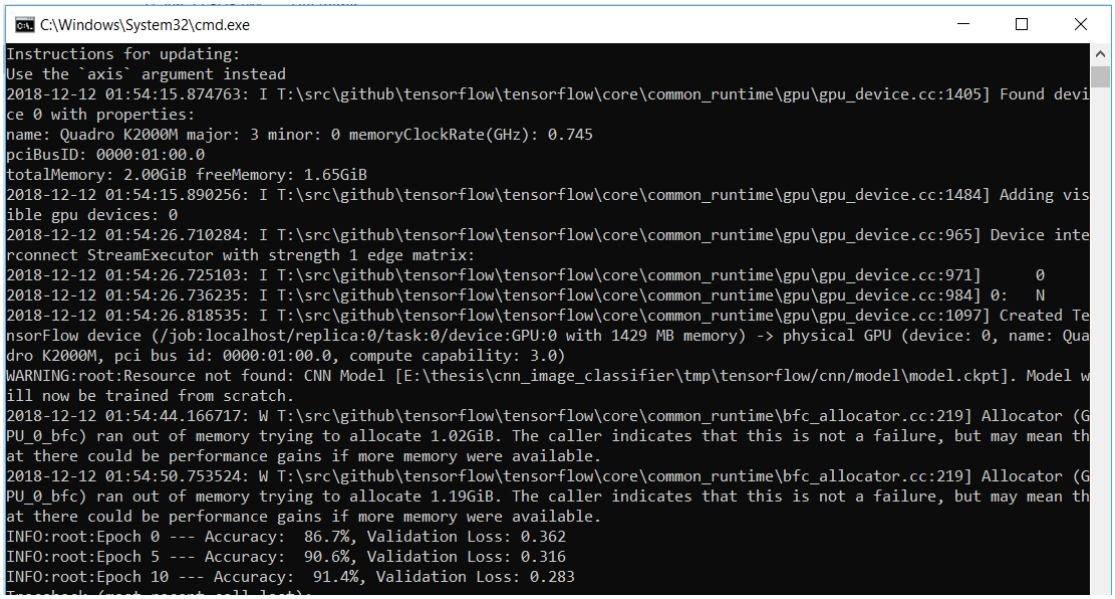


Figure 4.5: Training AlexNet

- **Prediction and Plotting**

  After the training is done, we can use the saved model weights to predict cancer type from the unknown test dataset. Our model predicts the outcome and plot the best match probability into input unknown image.

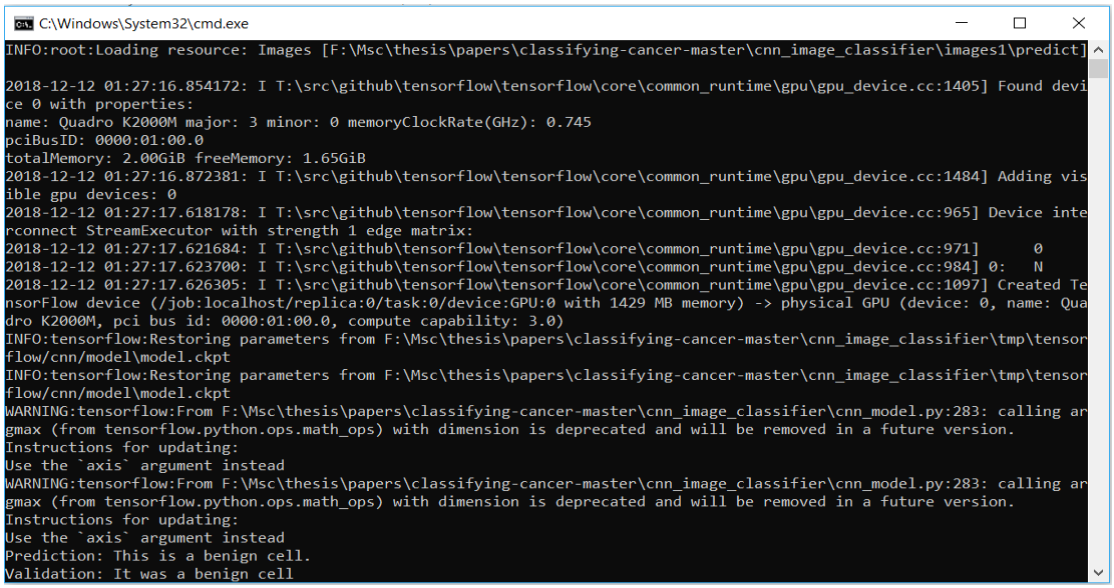


Figure 4.6: Predicting the Cancer Type

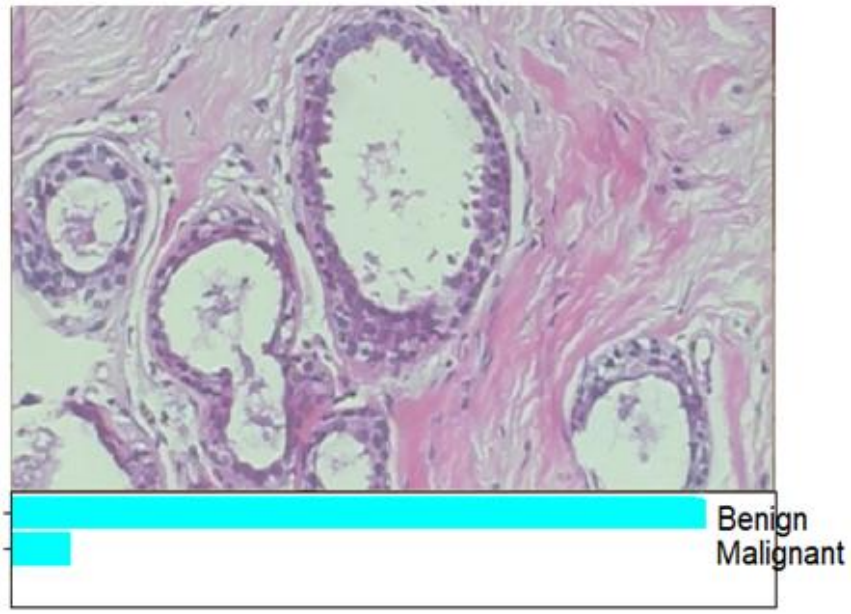


Figure 4.7: Plotted Model Outcome

CHAPTER 5

EXPERIMENTAL RESULT

In our proposed method we have used CNN with AlexNet instead of hand-crafted model. Hand crafted models give rise to unwanted complexity and might have inferior outcome than already existed benchmarked one.

Figure 13 shows the overall accuracy of the proposed agent over 2500 plus steps.

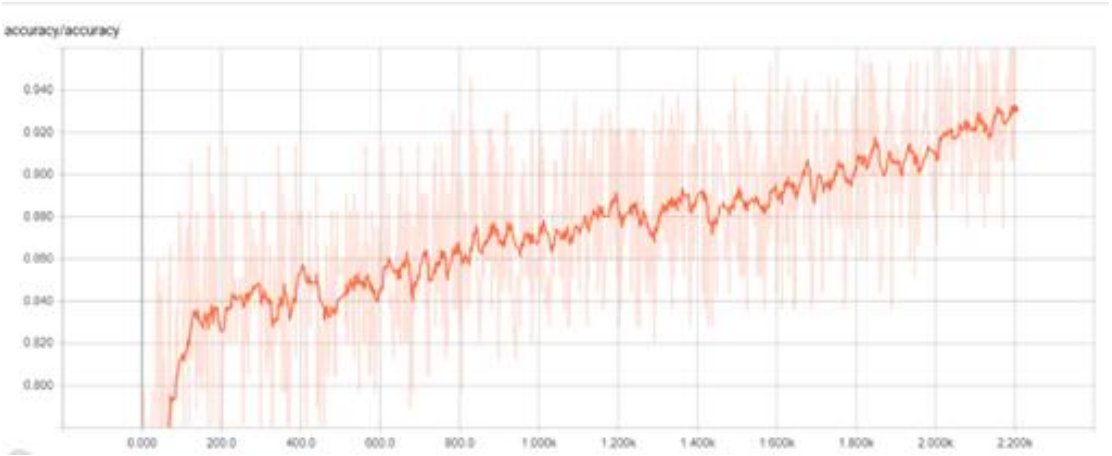


Fig 5.1: Overall accuracy of the agent

From above graph it's evident that the agent hits local maxima and local minima throughout the 2500 steps. To reduce the overfitting and being stuck in local maxima & minima we need to adjust the weight in each back-propagation step. Weight distribution in the connected layers can be seen from the figure 14.

Fig 5.2: Weight distribution in FC layer

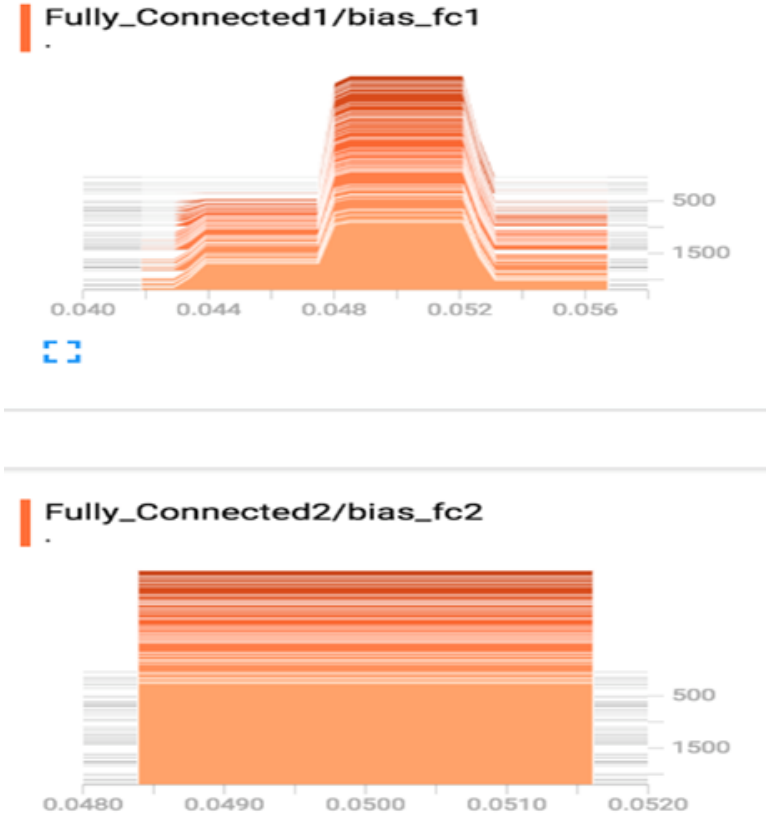


Fig 5.3: Bias distribution in FC layer

Along with the weights we need to adjust the bias in the fully connected layers. Bias distribution can be visualized from the figure 15

At first the agent learns at a steady pace and has cost function of ex at the beginning of the learning. Cost function of the agent can be visualized from the figure 8

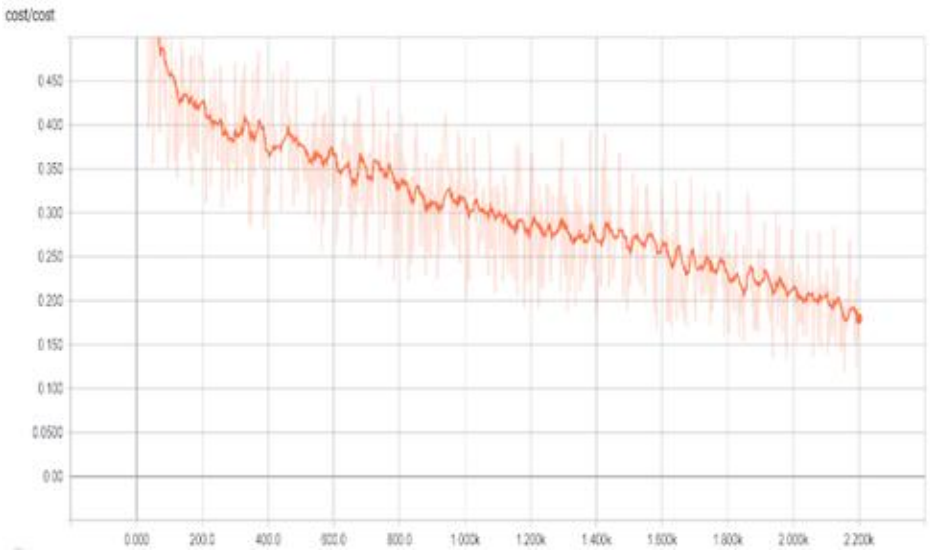


Fig 5.4: Cost function summery

the recognition rate of the agent is computed at the image level thus providing a means to estimate solely the image classification accuracy of the CNN models. Let

Iall be the number of cancer images of the test set. If the system classifies correctly Iclsfy cancer images, then the recognition rate at the image level is:

$$Image\ Recognition\ Rate = \frac{I_{all}}{I_{clsfy}} \times 100$$

Image recognition rate of the agent can be visualized from the table 2

Table 2 Agent Recognition Rate

Test Case	Magnification Factor			
	40x	100x	200x	400x
1	97.2%	95.7%	95.6%	89.8%
2	96.8%	94.1%	95.0%	90.0%
3	96.9%	93.0%	90.8%	89.7%

In each step 10 images from each magnification factor is considered. Image recognition rate can also be viewed from the following chart

Fig 5.5: Image Recognition Summery

The trained agent performs outstanding for 40x magnification factor. As the 400x magnified images has the most complex textures, agent's performance is decreased dramatically while classifying those images. Over time 400x images causes the agent to overfit and stuck in local maxima and minima.

Our proposed agent out performs the previous works at 40x and 100x magnification factor. But in 200x and 400x magnification image our agent has slightly less accuracy. It's due to the factor of hand-crafted models in the previous works. A details comparison is shown in the bellow chart

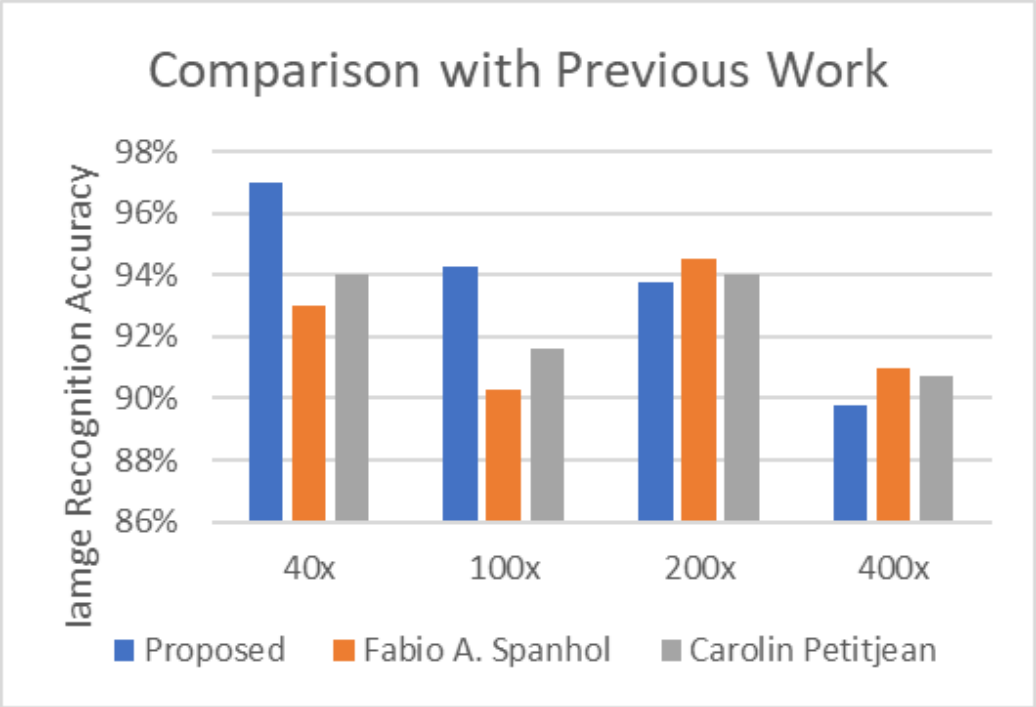


Fig 5.6: Comparison with Previous Work

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

We have presented am method to classify images from the BreakHis dataset to classify breast cancer type based on deep neural network. Our work has shown that existing models particularly Alexnet perform better than hand crafted models when classifying color images and object with diverse image patterns.

In our work we have used sliding window mechanism, that allow to deal with the high-resolution of textured images which can also perform as expected with low resolution images without changing the whole architecture. Experimental results have shown that our proposed method works better than exiting classifiers while classifying breast cancer type from BreakHis dataset zoom factor of 40x and 100x.

6.2 Future Work

Future work can explore different CNN architectures and the optimization of the hyperparameters. Our proposed method has slight error rate than then existing method while classifying 200x and 400x zoomed images. By exploring other pretrained models like Inception V3, GoogleNet etc would be a great option to get better accuracy.

Reference:

[1] Wikipedia, "Eventbrite," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Eventbrite. [Accessed 03 April 2018].

FinancesOnline.

[2] S. R. Lakhani, E. I.O., S. Schnitt, P. Tan, and M. van de Vijver, WHO classification of tumours of the breast, 4th ed. Lyon: WHO Press, 2012.

[3] B. Stenkvist, S. Westman-Naeser, J. Holmquist, B. Nordin, E. Bengtsson, J. Vegelius, O. Eriksson, and C. H. Fox, "Computerized nuclear morphometry as an objective method for characterizing human cancer cell populations," Cancer Research, vol. 38, no. 12, pp. 4688–4697, 1978.

[4] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, and R. Monczak, "Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images," Computers in Biology and Medicine, vol. 43, no. 10, pp. 1563–1572, 2013.

[5] P. Filipczuk, T. Fevens, A. Krzyz˙ak, and R. Monczak, "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies," IEEE Transactions on Medical Imaging, vol. 32, no. 12, pp. 2169–2178, 2013.

[6] Y. M. George, H. L. Zayed, M. I. Roushdy, and B. M. Elbagoury, "Remote computer-aided breast cancer detection and diagnosis system based on cytological images," IEEE Systems Journal, vol. 8, no. 3, pp. 949–964, 2014.

[7] Y. Zhang, B. Zhang, F. Coenen, and W. Lu, "Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles," Machine Vision and Applications, vol. 24, no. 7, pp. 1405– 1420, 2013.

[8] Y. Zhang, B. Zhang, F. Coenen, J. Xiau, and W. Lu, "One-class kernel subspace ensemble for medical image classification," EURASIP Journal on Advances in Signal Processing, vol. 2014, no. 17, pp. 1–13, 2014.

[9] S. Doyle, S. Agner, A. Madabhushi, M. Feldman, and J. Tomaszewski, "Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features," in Proceedings of the 5th IEEE International Symposium on Biomedical Imageging (ISBI): From Nano to Macro, vol. 61. IEEE, May 2008, pp. 496–499.

[10] A. J. Evans, E. A. Krupinski, R. S. Weinstein, and L. Pantanowitz, “2014 american telemedicine association clinical guidelines for telepathology: Another important step in support of increased adoption of telepathology for patient care,” Journal of Pathology Informatics, vol. 6, 2015.

[11] F. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, “A dataset for breast cancer histopathological image classification,” IEEE Transactions of Biomedical Engineering, 2016.

[12] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, pp. 1798–1828, 2013.

[13] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, 2015.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in Neural Infor- mation Processing Systems 25, 2012, pp. 1097–1105.

[15] X. X. Niu and C. Y. Suen, “A novel hybrid cnn–svm classifier for recognizing handwritten digits,” Pattern Recognition, vol. 45, no. 1318- 1325, 2012.

[16] L. G. Hafemann, L. S. Oliveira, and P. Cavalin, “Forest species recognition using deep convolutional neural networks,” in International Conference on Pattern Recognition, 2014, pp. 1103–1107.

[17] T. H. Vu, H. S. Mousavi, V. Monga, U. A. Rao, and G. Rao, “Dfdl: Discriminative feature-oriented dictionary learning for histopathological image classification,” in Proceedings of the IEEE 12th International Symposium on Biomedical Imaging (ISBI). IEEE, Apr. 2015, pp. 990– 994.

[18] A. L. Mescher, Junqueira’s basic histology: text and atlas. New York: McGraw-Hill Lange, 2013.

[19] F. Xing, Y. Xie, and L. Yang, “An automatic learning-based framework for robust nucleus segmentation,” IEEE Transactions on Biomedical Imaging, 2015.

[20] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, “Deep feature learning for knee cartilage segmentation using a tripla- nar convolutional neural network,” in Medical Image Computing and Computer-

Assisted Intervention – MICCAI 2013, ser. Lecture Notes in Computer Science, K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds. Springer Berlin Heidelberg, 2013, vol. 8150, pp. 246– 253.

[21] A. Cruz-Roa, J. Arevalo Ovalle, A. Madabhushi, and F. A. González Osorio, "A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, ser. Lecture Notes in Computer Science.

[22] K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds. Springer Berlin Heidelberg, 2013, vol. 8150, pp. 403–410.

[23] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in Proceedings of 26th Annual Conference on Neural Information Processing Systems 2012 (NIPS).

[24] P. L. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Dec. 2012, pp. 2852–2860. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.

[25] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS). Springer-Verlag, Jun. 2010, pp. 253–256.

[26] D. Scherer, A. Müller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in Proceedings of the 20th International Conference on Artificial Neural Networks: Part III.

[27] K. Diamantaras, W. D. Duch, and L. S. Iliadis, Eds. Springer- Verlag, Sep. 2010, pp. 92–101.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Proceedings of 26th Annual Conference on Neural Information Processing Systems 2012 (NIPS).

[29] P. L. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Dec. 2012, pp. 1106–1114. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-

convolutional-neural-networks.

[30] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, pp. 1915–1929, 2013.

[31] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, pp. 2278–2324, 1998.

[32] L. Bottou, "Stochastic gradient tricks," in Neural Networks, Tricks of the Trade, Reloaded, ser. Lecture Notes in Computer Science (LNCS 7700), G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Springer, 2012, pp. 430–445. [Online]. Available: http://leon.bottou.org/papers/bottou-tricks-2012

[33] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 3, pp. 226–239, 1998.

[34] NVIDIA Corporation. (2015) Nvidia tesla product literature. [Online].Available http://www.nvidia.com/object/tesla_product_literature.html

[35] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," arXiv preprint arXiv:1408.5093, 2014.

[36] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," Biological Cybernetics, vol. 61, pp. 103–113, 1989.

[37] C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, pp. 55–73, 1990.

[38] N. A. Hamilton, R. S. Pantelic, K. Hanson, and R. D. Teasdale, "Fast automated cell phenotype image classification," BMC Bioinformatics, vol. 8, 2007. [Online]. Available: http://www.biomedcentral.com/ 1471-2105/8/110

[39] L. P. Coelho, A. Ahmed, A. Arnold, J. Kangas, A. S.Sheikh, E. P. Xing, W. Cohen, and R. F. Murphy, "Structured literature image finder: extracting information from text and images in biomedical literature," in Linking Literature, Information, and Knowledge for Biology, ser. LNCS,

[40] P. L. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Advances in Neural Information Processing Systems

25: 26th Annual Conference on Neural Information Processing Systems 2012., Dec. 2012. [Online]. Available: http://papers.nips.cc/ book/advances-in-neural-information-processing-systems-25-2012.