

APPROVAL OF BOARD OF EXAMINERS

The thesis titled “DETECTION OF BRAIN TUMORS FROM MRI IMAGES USING MACHINE LEARNING APPROACHES” submitted by Mohammad Shahjahan Majib, Roll No: 0418140002, Session: 2017 - 2018 has been accepted as satisfactory in partial fulfillment of the requirements for the degree of Master of Science in Computer Science and Engineering on 01 February 2021.

BOARD OF EXAMINERS

1. _____
Name: Dr Md. Mahbubur Rahman
Designation: Professor
Affiliation: Department of CSE, MIST, Dhaka
Chairman
(Supervisor)

2. _____
Name: Brig Gen A B M Humayun Kabir
Designation: Senior Instructor and Head of the Department
Affiliation: Department of CSE, MIST, Dhaka
Member
(Ex Officio)

3. _____
Name: Dr. T. M. Shahriar Sazzad
Designation: Assistant Professor
Affiliation: Department of CSE, MIST, Dhaka
Member
(Internal)

4. _____
Name: Dr. Md. Shohrab Hossain
Designation: Professor
Affiliation: Department of CSE, BUET, Dhaka
Member
(External)

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. The thesis (fully/ partially) has not been submitted for any degree or diploma in any university or institute previously.

Mohammad Shahjahan Majib

Department of Computer Science and Engineering
Military Institute of Science and Technology
01 February 2021

SUMMARY

Tumor indicates unfettered presence of a cluster of cells in a specific area of the body part. Brain tumor is considered one of the most common tumors for both men and women and can lead to high death risk if patients fail to obtain appropriate medical treatment. In order to diagnose brain tumors, electronic modalities are integrated and among them MRI is a popular one. For MRI brain tumor region analysis segmentation, detection and classification are considered as important steps in digital imaging pathology laboratory. Existing state-of-the-art approaches demand widespread amount of supervised training data from pathologists and may still accomplish poor results in images from unseen tissue types. In this study, two approaches each for Conventional ML and Deep ML have been presented to identify brain tumors from MRI images. For Conventional ML, a suitable framework has been presented where for the first time in compare to all other existing accessible approaches morphological operations has been incorporated to eliminate undesirable regions and to assist segmentation and identification of region of interests. Compared with existing state-of the-art supervised models, our method generalizes considerably improved identified results on brain tumor cells deprived of training data. Even with training data, our approach attains the identical performance without supervision cost. This study results indicates an accuracy rate above 96.23% accuracy associated to existing works. For Deep ML, another framework is presented where 97.84% accuracy is achieved.

ACKNOWLEDGEMENT

The author would like to express his sincere gratitude to Professor Dr Md. Mahbubur Rahman of Department of Computer Science and Engineering, Military Institute of Science and Technology, Mirpur Cantonment for his kind supervision, guidance, encouragement and inspection throughout the entire research period. The author also recognizes Assistant Professor Dr. T. M. Shahriar Sazzad of Department of Computer Science and Engineering, Military Institute of Science and Technology, Mirpur Cantonment for his valuable guidance. At the same time, the author expresses thankful gratitude to Dr. Md. Sadequel Islam Talukder for supporting the work while providing all the necessary MRI brain tumor test images along with feature information to identify the ROIs. This study also thanks Military Institute of Science and Technology for their support to carry out this study successfully.

TABLE OF CONTENTS

SUMMARY.....	iii
LIST OF FIGURES.....	vii
LIST OF TABLES.....	ix
CHAPTER 1: INTRODUCTION.....	1
1.1 Background.....	1
1.2 Problem Statement.....	2
1.3 Thesis Objectives.....	2
1.4 Methodology.....	3
1.5 Thesis Scope.....	4
1.6 Organization of the Thesis.....	5
CHAPTER 2: LITERATURE REVIEW.....	6
2.1 Medical Imagery.....	6
2.1.1 Existing Available Approaches to Analyze Tissues.....	7
2.1.2 Traditional Microscopic Analysis.....	7
2.1.3 Capturing the Images Using Electronic Devices (e.g. Ultrasound, CT, MRI, PET, X-Ray) and Analysis.....	8
2.1.4 Computer Based Approaches Using Digitized Images Acquired from the Slides of Biopsy Specimen.....	9
2.2 Image Processing Techniques.....	10
2.2.1 Pre-processing.....	10
2.2.1.1 Enhance Contrast.....	10
2.2.1.2 Noise Removal Approach.....	11
2.2.2 Segmentation.....	12
2.2.3 Identification and/Classification.....	13
2.3 Discussion.....	14
2.4 Summary.....	20
CHAPTER 3: METHODOLOGY.....	21
3.1 Introduction.....	21
3.2 Research Methodology.....	21
3.3 Research Approach.....	21
3.4 Research Activity.....	22
3.4.1 Data Collection.....	22
3.4.2 Development of Proposed Approach.....	23
3.4.3 Dataset Used in this Study.....	24
3.4.4 Validation of the Proposed Approach.....	24
3.5 Techniques.....	26
3.5.1 Pre-processing.....	26
3.5.1.1 Intensity Enhancement.....	27
3.5.1.2 Elimination of Unwanted Regions for Image Analysis.....	27
3.5.1.2.1 Mean Filter.....	27
3.5.1.2.2 Median Filter.....	28
3.5.1.2.3 Filter using Mathematical Morphological Operation.....	29
3.5.2 Segmentation.....	29
3.5.2.1 Threshold Based Segmentation.....	29
3.5.2.2 Clustering Approach.....	31
3.5.2.2.1 K-means Clustering Approach.....	31
3.5.2.2.2 Feature Extraction.....	32
3.5.3 Identification and/ Classification.....	33

3.5.3.1 Identification.....	33
3.5.3.2 Classification.....	34
3.5.3.2.1 SVM.....	34
3.5.3.2.2 Naïve Bayes (NB).....	37
3.5.3.2.3 Convolutional Neural Network (CNN).....	38
3.5.3.2.4 Decision Tree (DT).....	39
3.6 Data Analysis Techniques.....	39
3.6.1 Selection Criteria of Analysis.....	40
3.6.2 Evaluation of Results.....	40
3.7 Summary.....	40
CHAPTER 4: EXPERIMENTAL RESULTS.....	41
4.1 Research Activity 1.....	41
4.1.1 Research Approach One.....	41
4.1.2 Research Approach Two.....	43
4.1.3 Experiment Results.....	45
4.1.4 Summary.....	57
4.2 Research Activity 2.....	58
4.2.1 Effects of Filter Size.....	58
4.2.2 Effects of Threshold Function.....	63
4.2.3 Selection of Pool Size.....	66
4.2.4 Performance of Classifiers.....	67
4.2.5 Summary.....	80
CHAPTER 5: CONCLUSION.....	81
5.1 Main Outcomes.....	81
5.1.1 Research Activity One.....	81
5.1.2 Research Activity Two.....	82
5.2 Thesis Contributions.....	83
5.3 Thesis Limitations.....	84
5.4 Future Works.....	85
References.....	86

LIST OF FIGURES

Figure: 3-1 Research Approach.....	22
Figure 3-2: Research activity plan (on the right), with steps for the proposed approach (on the left).....	23
Figure 3-3: an example of expert marked (b) and test (a) images which are identical to each other collected from the medical laboratory and used in this research study.....	25
Figure 3-4: Kernel function and mapping.....	35
Figure 3-5: Kernel machine.....	37
Figure 3-6: A simple Neural Network and Convolutional Neural Network.....	36
Figure 4-1: Detailed flow-chart for research approach one.....	42
Figure 4-2: Detailed flow-chart for research approach two.....	43
Figure 4-3: Filter operation using Gaussian filter, Median filter and Morphological Operation (test image-1).....	45
Figure 4-4: Filter operation using Gaussian filter, Median filter and Morphological Operation (test image-2).....	46
Figure 4-5: Filter operation using Gaussian filter, Median filter and Morphological Operation (test image-3).....	47
Figure 4-6: Filter operation using Gaussian filter, Median filter and Morphological Operation (test image-4).....	48
Figure 4-7: Filtered image using different filters (test image-1).....	49
Figure 4-8: Filtered image using different filters (test image-2).....	50
Figure 4-9: Filtered image using different filters (test image-3).....	51
Figure 4-10: Filtered image using different filters (test image-4).....	52
Figure 4-11: Identification of results for research approach one and research approach two (test image- 1).....	53
Figure 4-12: Identification of results for research approach one and research approach two (test image- 2).....	54
Figure 4-13: Identification of results for research approach one and research approach two (test image- 3).....	55
Figure 4-14: Identification of results for research approach one and research approach two (test image- 4).....	56
Fig 4-15: Confusion Matrix for selecting filter size 40.....	59
Fig 4-16: Train vs Validation Accuracy for each Epoch for filter size 40.....	59
Fig 4-17: Train vs Validation Loss for each Epoch for filter size 40.....	59
Fig 4-18: Confusion Matrix for selecting filter size 28.....	60
Fig 4-19: Train vs Validation Accuracy for each Epoch for filter size 28.....	60
Fig 4-20: Train vs Validation Loss for each Epoch for filter size 28.....	60
Fig 4-21: Confusion Matrix for selecting filter size 16.....	61
Fig 4-22: Train vs Validation Accuracy for each Epoch for filter size 16.....	61
Fig 4-23: Train vs Validation Loss for each Epoch for filter size 16.....	61
Fig 4-24: Confusion Matrix for selecting filter size 32.....	62
Fig 4-25: Receiver Operating Characteristic (ROC) Curve for CNN.....	62
Fig 4-26: Train vs Validation Accuracy for each Epoch for filter size 32.....	62
Fig 4-27: Train vs Validation Loss for each Epoch for filter size 32.....	62
Fig 4-28: Confusion Matrix after selecting ReLU activation function in dense layer...	63
Fig 4-29: Train vs Validation Accuracy for each Epoch after selecting ReLU activation function in dense layer.....	63
Fig 4-30: Train vs Validation Loss for each Epoch after selecting ReLU activation	

function in dense layer.....	63
Fig 4-31: Confusion matrix after selecting softmax as activation function in dense layer.....	64
Fig 4-32: Train vs Validation Accuracy for each Epoch after selecting softmax activation function in dense layer.....	64
Fig 4-33: Train vs Validation Loss for each Epoch after selecting softmax activation function in dense layer.....	64
Fig 4-34: Confusion matrix after ReLU as activation function in second layer.....	65
Fig 4-35: Train vs Validation Accuracy for each Epoch after ReLU as activation function in second layer.....	65
Fig 4-36: Train vs Validation Loss for each Epoch after ReLU as activation function in second layer.....	65
Fig 4-37: Confusion matrix after 2x2 maxpooling.....	66
Fig 4-38: Train vs Validation Accuracy for each Epoch after 2x2 maxpooling.....	66
Fig 4-39: Train vs Validation Loss for each Epoch after 2x2 maxpooling.....	66
Fig 4-40: Receiver Operating Characteristic (ROC) Curve for SVM.....	68
Fig 4-41: Confusion Matrix for SVM.....	68
Fig 4-42: Receiver Operating Characteristic (ROC) Curve for RF.....	71
Fig 4-43: Confusion Matrix for RF.....	71
Fig 4-44: Receiver Operating Characteristic (ROC) Curve for DT.....	73
Fig 4-45: Confusion Matrix for DT.....	73
Fig 4-46: Receiver Operating Characteristic (ROC) Curve for KNN.....	75
Fig 4-47: Confusion Matrix for KNN.....	75
Fig 4-48: Receiver Operating Characteristic (ROC) Curve for MLP.....	78
Fig 4-49: Confusion Matrix for MLP.....	78
Fig 4-50: Receiver Operating Characteristic (ROC) Curve for NB.....	79
Fig 4-51: Confusion Matrix for NB.....	79

LIST OF TABLES

Table 2.1: Medical image analysis on MRI brain tumor images (Conventional ML)...	16
Table 2.2: Medical image analysis on MRI brain tumor images (Deep ML).....	18
Table 2.3: Important Literature findings.....	19
Table 3-1: Composition of research test image dataset.....	24
Table 3-2: Partition of the full datasets into training dataset and unseen test dataset...	25
Table 4-1: Performance of different models.....	80