



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

بسم الله الرحمن الرحيم



HANAA ALY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم



شبكة المعلومات الجامعية التوثيق الإلكتروني والميكروفيلم



HANAA ALY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات



يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار



HANAA ALY



Information Systems Department
Faculty of Computer and Information Sciences
Ain Shams University

Real Time Implementation of Medical Data Analysis Approaches

A thesis submitted to the Department of Information Sciences, Faculty
of Computers and Information Sciences, Ain Shams University, in partial
fulfillment of the requirements for the Degree of Doctor of Philosophy in Computer and Information Sciences.

By:
Yasmeen Farouk Bakry

Supervisors:

Prof. Dr. Hossam ElDeen Faheem

Professor at Computer Systems Department
Faculty of Computer and Information Sciences
Ain Shams University

Dr. Sherine Rady

Associate Professor at Information Systems Department
Faculty of Computer and Information Sciences
Ain Shams University

February 2021

Dedicated to
my beloved mother
and the memory of my father who lost his life fighting Alzheimer's
May this work help others in their battle...

Acknowledgements

First and foremost, all praise is due to Allah, we worship Him, we seek His help and His forgiveness. The Prophet Muhammad (pbuh) said: “He who does not thank the people is not thankful to Allah.” [Sunan Abi Dawud, graded Sahih by Shaikh al-Albani.]

Hence, I would like to express my sincere gratitude to my supervisors Prof. Dr. Hossam M. Faheem and Associate Prof. Dr. Sherine Rady for their guidance, support, encouragement, and mostly their patience. Their trust in me made them more of mentors than professors. I doubt that I will ever be able to convey my appreciation but I owe them my eternal respect and gratitude.

I owe my deepest gratitude to my steadfast husband, Tarek Eldeeb for his inexpressible patience and dedicated support throughout my research. He saw me through and shared with me my wish to complete this research successfully despite all the challenges we faced. His sincere belief in me made all of this possible. I also wish to thank my lovely kids for enduring my absence with their love and care. They are always a true source of inspiration.

Finally, I am forever grateful for the continuous support and encouragement of my loving mother, my brothers Ahmad and Hazem, and my sister Shereen, not only throughout my research but through the journey of life as well. Their love and support have made me who I am.

Abstract

With the increasing size of medical datasets and their high dimensionality nature there arises a significant need for methods to accelerate the processing and analysis. High dimensional medical images may consume several hours and even days executing exhaustive analysis processes, such as those employed in the classification of Alzheimer's disease.

Alzheimer's disease is a serious form of dementia. In 2018, fifty million people worldwide has been reported living with dementia. This number will reach 152 million people by 2050 [95]. This vastly expanding disease causes a progressive decline in the mental abilities that usually starts with memory loss and ends with cognitive and behavioral disorders. With no current cure, treatments focus on slowing the progression of the disease and controlling its symptoms. Early diagnosis by studying structural MRI (sMRI) is the key.

Image registration of sMRI is a serious challenge in the field of medical image analysis generally and for Alzheimer's disease analysis specifically. The correct localization of brain tissue deformation and determination of the tumor growth rely majorly on the accuracy of the image registration process. Image registration techniques are computationally intensive time-consuming tasks. Field Programmable Gate Arrays (FPGAs) have fast-evolving and customizable hardware acceleration capabilities that promise to help speed up these tasks.

The work in this thesis presents analytical predictive solutions for the classification of Alzheimer's disease problem using unsupervised and supervised learning techniques. The supervised learning solution run in an accelerated framework of a software/hardware co-design model using FPGA hardware chips.

The proposed image analysis uses texture features extracted from the gray level co-occurrence matrix and voxel-based morphometry (VBM) neuroimaging analysis to classify Alzheimer's disease patients. VBM initially explores the differences in brain anatomy using statistical parametric mapping. Unsupervised clustering techniques are investigated for an early diagnosis of Alzheimer's disease. Different supervised classification techniques are studied as well to come up with the best classification accuracy. The studied techniques are support vector machine, k-nearest neighbor, and decision tree. Moreover, feature selection using entropy and feature reduction using principle component analysis are employed with supervised learning techniques to overcome the curse of dimensionality hence contributing to faster response. The thesis explores as well the discriminative power of certain anatomical regions of interest (ROI) for classifying Alzheimer's disease.

The proposed classification technique is accelerated using an optimized software/hardware co-design model where a part of the VBM analysis runs on CPU, while the rest runs on FPGA. FPGA's acceleration power is used in the implementation of the diffeomorphic image registration algorithm DARTEL as a part of the registration process in the VBM analysis. An optimized and pipelined hardware architecture is proposed.

The proposed clustering technique provided a preliminary insight with an accuracy 76%, while the proposed classification technique managed successfully to differentiate between Alzheimer's disease patients and normal controls with an accuracy of 93%. The proposed classification technique proved as well that analyzing brain regions that include Hippocampus, Cerebelum Left, Cerebelum Right, and Calcarine would lead to better time performance results than analyzing the whole brain regions. This set of ROI discriminators recorded 4% increase in the accuracy compared to whole brain method with only 7% of the whole brain feature vector and a 16x speedup in the run time. The hardware acceleration of the DARTEL registration algorithm resulted in a maximum speed up factor of 114x on function-level, compared to the CPU, with a contribution of 8x faster for the overall performance in the registration process of the software tool.

Table of contents

List of figures	xiii
List of tables	xvii
Nomenclature	xix
1 Introduction	1
1.1 Motivation	1
1.2 Problem Statement	3
1.3 Research Objectives	4
1.4 Research Contributions	4
1.5 Thesis Outlines	6
2 Scientific Background	9
2.1 Alzheimer’s Disease	9
2.2 Neuroimaging Techniques	13
2.2.1 Magnetic Resonance Imaging	14
2.2.2 Functional Magnetic Resonance Imaging	16
2.2.3 Positron Emission Tomography	17
2.2.4 Single Photon Emission Computed Tomography	18
2.3 Machine Learning Techniques	19
2.3.1 Gray Level Co-occurrence Matrix	19

2.3.2	K-means	20
2.3.3	K-medoids	21
2.3.4	Support Vector Machine	21
2.3.5	K-Nearest Neighbour	23
2.3.6	Decision Tree	25
2.3.7	Principle Component Analysis	27
2.3.8	Entropy	28
2.4	DARTEL Image Registration Algorithm	28
2.5	Hardware Acceleration	31
2.5.1	Field Programmable Gate Arrays	32
2.5.2	Graphics Processing Units	35
2.6	Summary	36
3	Related Work	37
3.1	Neuroimaging Analysis	37
3.2	Intelligent Analytical Techniques	41
3.2.1	Clustering	41
3.2.2	Classification	42
3.2.3	Texture Analysis	45
3.2.4	Feature Reduction	46
3.2.5	Image Registration	47
3.3	Hardware Acceleration in Medical Image Analysis	48
3.3.1	FPGA	48
3.3.2	GPU	49
3.4	Summary	51
4	Unsupervised Clustering Approach for Early AD Diagnosis	53
4.1	Framework of the Proposed Unsupervised Clustering Approach	53

4.2	MRI Pre-Processing	54
4.3	Voxel-Based Morphometry Analysis	58
4.3.1	Segmentation	58
4.3.2	Registration	58
4.3.3	Normalization and Smoothing	60
4.3.4	Statistical Analysis	61
4.4	Region of Interest Masking	62
4.5	Experiments and Results	62
4.5.1	Dataset	62
4.5.2	Experimental Setup and Evaluation Measures	64
4.5.3	Performance Analysis of k-means and k-medoids on Whole-Brain	64
4.5.4	Performance Analysis of k-means and k-medoids on ROIs	65
4.6	Summary and Discussion	68
5	Supervised learning approach for AD Identification	69
5.1	Framework of the Proposed Supervised AD Identification	69
5.2	Voxel Morphometrics Extraction using VBM	70
5.3	Texture Extraction using GLCM	70
5.4	Feature Selection using Entropy	71
5.5	Extraction of Significant ROIs	71
5.6	Feature Reduction using PCA	72
5.7	Classification	72
5.8	Experiments and Results	73
5.8.1	Dataset	73
5.8.2	Experimental Setup and Evaluation Measures	73
5.8.3	Performance Analysis of VBM-GLCM Features Combination	74
5.8.4	Performance Analysis of Entropy-based Feature Selection	76
5.8.5	Performance Analysis of PCA Feature reduction of GLCM Features	78

5.8.6	Performance Analysis of Whole Brain Versus ROI-Based Approach	79
5.8.7	Performance Analysis of Machine Learning Classifiers	83
5.8.8	Performance Analysis of PCA on Machine Learning Classification .	92
5.9	Summary and Discussion	94
6	Accelerated Software/Hardware Co-Design Model	95
6.1	Profiling VBM Run Time	95
6.2	Proposed MRI Registration Software/Hardware co-design Model	97
6.3	Software/Hardware Co-design Interfacing	99
6.4	FPGA Design Optimization	100
6.4.1	Fixed-Point Data Type	100
6.4.2	Complex and Non-linear Arithmetic Functions	101
6.4.3	Loop unrolling and function in-lining	102
6.4.4	Pipelining	102
6.5	Experiments and Results	103
6.5.1	Dataset	103
6.5.2	Experimental Setup and Evaluation Measures	103
6.5.3	Performance Analysis of FPGA implementation	103
6.6	Summary and Discussion	111
7	Conclusion and Future Work	115
7.1	Summary	115
7.2	Directions for Future Work	118
	References	121

List of figures

1.1	Number of people with dementia (in millions) [10]	2
1.2	Percentage of Increase in some Causes of Death Between the year 2000 and 2019 [17]	3
2.1	AD Continuum [10]	10
2.2	Difference between AD patient brain and NC brain. [7]	11
2.3	Different Parts of the brain and their functions. [111]	12
2.4	CSF within Cerebellum Area: Normal and Patient Brains.	13
2.5	sMRI: Sagittal View.	15
2.6	T1-weighted MRI sequence: GM, WM, and CSF are highlighted.	16
2.7	A sample of an fMRI Image showing regions of brain activation.	17
2.8	A sample of PET Image.	18
2.9	A sample of a SPECT Image.	19
2.10	DARTEL Template Generation: Intensity averages of GM images (coronal, sagittal and axial views) after multiple iterations a) average after initial rigid-body alignment b) average after two iterations c) average after four iterations d) average after six iterations.	29
2.11	FPGA Structure and Interconnects. [68]	33
2.12	SPARTAN FPGA Board.	34
2.13	GPU Card.	35