



FINAL YEAR PROJECT REPORT

**AUTOMATED CLASSIFICATION OF
CERVICAL CELL IMAGES
FOR CANCER SCREENING**

**In fulfillment of the requirement
For degree of
BS (COMPUTER SCIENCES)**

By

**FAISAL JAMIL AHMED
ARSALAN AHMED
OSAID FAISAL**

**70155 (BSCS)
67741 (BSCS)
70150 (BSCS)**

SUPERVISED

BY

DR. RAHEEL SIDDIQI

BAHRIA UNIVERSITY (KARACHI CAMPUS)

SPRING-2024

DECLARATION

We hereby declare that this project report is based on our original work except for citations and quotations which have been duly acknowledged. We also declare that it has not been previously and concurrently submitted for any other degree or award at Bahria University or other institutions.

Signature : Faisal

Name : Faisal Jamil Ahmed

Reg No. : 70155

Signature : Arsalan

Name : Arsalan Ahmed

Reg No. : 67741

Signature : Osaid

Name : Osaid Faisal

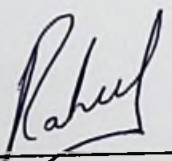
Reg No. : 70150

Date : 12-06-2024

APPROVAL FOR SUBMISSION

We certify that this project report entitled **“AUTOMATED CLASSIFICATION OF CERVICAL CELL IMAGES FOR CANCER SCREENING”** was prepared by **FAISAL JAMIL AHMED, ARSALAN AHMED, & OSAID FAISAL** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Science in **Computer Science** at Bahria University, Karachi.

Approved by,

Signature :  _____

Supervisor: **Dr. Raheel Siddiqi**

Date : 13-06-2024

The copyright of this report belongs to Bahria University according to the Intellectual Property Policy of Bahria University BUORIC-P15 amended on April 2019. Due acknowledgement shall always be made of the use of any material contained in, or derived from, this report.

© 2024 Bahria University. All right reserved.

Specially dedicated to -
My beloved grandmother, mother and father
(Fazal Fazal Ahmed)
My beloved grandmother, mother and father
(Ammar Ahmed)
My beloved grandmother, mother and father
(Chaudhary Fazal)

ACKNOWLEDGEMENTS

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express my gratitude to my project supervisor, Dr. Hadeel Baidag, for her invaluable advice, guidance, and her consistent patience throughout the development of this project.

Specially dedicated to

My beloved grandmother, mother and father

(Faisal Jamil Ahmed)

My beloved grandmother, mother and father

(Arsalan Ahmed)

My beloved grandmother, mother and father

(Osaid Faisal)

ACKNOWLEDGEMENTS

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express my gratitude to my project supervisor, **Dr. Raheel Siddiqi** for her invaluable advice, guidance, and her enormous patience throughout the development of the research.

In addition, we would also like to express our gratitude to our loving parents and friends who have helped and given us encouragement.

**AUTOMATED CLASSIFICATION OF CERVICAL CELL IMAGES
FOR CANCER SCREENING**

ABSTRACT

Cervical cancer remains a leading cause of mortality among women worldwide, with multiple deaths annually. Effective screening programs that facilitate early detection are critical for improving outcomes and saving lives. The current standard relies on manual examination of Pap smear slides by skilled cytotechnologists, which is time-consuming, costly, and prone to human error. This project proposes an automated cervical cancer screening system using deep learning and computer vision techniques to analyze cytology images. A convolutional neural network classifier will be developed to categorize cervical cells extracted from whole slide images into normal, and abnormal classes. The model will be trained on a dataset of Pap smear imagery with expert cell classifications. Extensive experiments will tune the architecture, hyper parameters, and regularization strategies to optimize performance. Advanced techniques including transfer learning and data augmentation will be employed to enhance accuracy and generalizability. Once validated, this artificial intelligence-powered screening assistant could greatly increase the speed, consistency, and accessibility of cervical cancer detection. By reliably flagging potentially cancerous cells, interventions could be administered sooner, preventing disease progression. The approach could be extended to improve outcomes for a variety of cancer types. This project will demonstrate the profound impact that AI and machine learning can have on life-saving medical diagnostics.

2.2	Automated Pap Smear Analysis	8
2.3	Image Analysis Technique	8
2.3.1	Pre-processing	8
2.3.2	Hand-Crafted Feature Extraction	8
2.3.3	Deep Learning Feature Representation	9
2.3.4	Classification	9
2.3.5	Ensemble Method	9
2.4	Classification Method	10
2.4.1	Traditional Machine Learning Models	10
2.4.2	Deep Convolutional Neural Network	10
2.4.3	Binary vs Multi Classification	10

Table of Contents

DECLARATION	ii
APPROVAL FOR SUBMISSION	iii
ACKNOWLEDGEMENTS	vi
ABSTRACT	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xiv
LIST OF SYMBOLS / ABBREVIATIONS	xvi
LIST OF APPENDICES	xvii
 INTRODUCTION	 1
1.1 Background	1
1.1.1 Herlev vs Sipakmed Dataset	2
1.1.2 Normal vs Abnormal Cell Difference	2
1.2 Problem Statement	3
1.3 Aims and Objectives	3
1.4 Scope of Project	3
 CHAPTER 2	 5
LITERATURE REVIEW	5
2.1 Introduction	5
2.2 Automated Pap Smear Analysis	6
2.3 Image Analysis Technique	8
2.3.1 Pre-processing	8
2.3.2 Hand-Crafted Feature Extraction	8
2.3.3 Deep Learning Feature Representation	9
2.3.4 Classification	9
2.3.5 Ensemble Method	9
2.4 Classification Method	10
2.4.1 Traditional Machine Learning Models	10
2.4.2 Deep Convolutional Neural Network	10
2.4.3 Binary vs Multi Classification	10

	2.4.4 Model Interpretability	11
2.5	Analysis	11
2.6	Convolutional Neural Network	12
2.7	Comprehensive Analysis and Future Directions	15
3	REQUIREMENT SPECIFICATIONS	17
3.1	Introduction	17
3.2	Existing System	17
3.3	Proposed System	18
3.4	Requirements Specification	19
	3.4.1 Functional Requirements	19
	3.4.2 Non-Functional Requirements	25
	3.4.3 User Interface Requirements	26
	3.4.4 Ethical and Legal Considerations:	28
3.5	Use Cases	28
	3.5.1 Doctor Use Cases:	29
	3.5.2 Patient Use Cases:	30
4	DESIGN AND METHODOLOGY	33
4.1	Data Acquisition and Pre-processing (For Binary Classification)	33
4.2	Model Architecture	34
4.3	CNN for Image Classification	35
4.4	Model Development	37
	4.4.1 Training Methodology	38
	4.4.2 Evaluation Metrics	39
	4.4.3 Confusion Matrix	39
	4.4.4 Accuracy	40
	4.4.5 Precision	40
	4.4.6 Recall	40
	4.4.7 F1-Score	41
4.5	Classification Performance	41
4.6	Dataset Pre-processing (For Multi Classification)	45

4.6.1	Folder Structure Creation	45
4.6.2	Image Resizing	45
4.6.3	Border Addition	45
4.6.4	Gaussian Blur	45
4.6.5	Image Format Conversion	46
4.7	Dataset Split	46
4.8	Data Augmentation	46
4.8.1	Random Zooming	46
4.8.2	Random Shearing	47
4.8.3	Horizontal Flipping	47
4.8.4	Rescaling	47
4.9	Model Architecture	47
4.9.1	Convolutional Neural Network	47
4.9.2	AlexNet	48
4.10	Ensemble Model	49
4.11	Training and Optimization	50
4.11.1	CNN Model	50
4.11.2	AlexNet Model	51
4.11.3	Ensemble Model	51
4.12	RoboFlow Integration	52
4.13	Ultralytics YOLO Model	52
4.14	Evaluation Metrics	52
4.14.1	Accuracy	53
4.14.2	Classification Report	53
4.14.3	Confusion Matrix	53
4.14.4	Model Contribution	54
CHAPTER 5		61
SYSTEM IMPLEMENTATION		61
5.1	Frontend Implementation	61
5.1.1	Authentication Module:	61
5.1.2	Sign In Module:	61
5.1.3	Sign Up Module:	61

5.1.4	Image Upload Module:	61
5.1.5	Report Generation Module:	62
5.1.6	Previous Records Module:	62
5.1.7	Logout Module:	62
5.1.8	Authentication Module:	62
5.1.9	Sign In Module:	62
5.1.10	Report Viewing Module:	63
5.2	Backend Implementation	63
5.2.1	RoboFlow API Integration	63
5.2.2	Supabase Database Management	63
5.3	Machine Learning Model	65
5.3.1	Model Integration	65
5.4	Flow:	65
5.4.1	Doctor WorkFlow	65
5.4.2	Patient Workflow	65
5.5	Implementation Consideration	65
5.5.1	Error Handling	65
5.5.2	Security	66
5.5.3	Scalability	66
5.5.4	Performance	66
5.5.5	Testing	66
5.6	Conclusion	66
CHAPTER 6		68
SYSTEM TESTING AND EVALUATION		68
6.1	Introduction	68
6.2	Testing Process Overview	68
6.2.1	Significance of Testing in Software Development	68
6.3	Test Planning	69
6.4	Test Cases	71
6.5	Test Approaches and Techniques	73
6.5.1	Automated Testing	73
6.5.2	Load Testing	73
6.5.3	Regression Testing	74

		xii
	6.5.4 Usability Testing	74
	6.5.5 Unit Testing	74
	6.5.6 Integration Testing	74
	6.5.7 End-to-End Testing	75
	6.5.8 User Acceptance Testing (UAT)	75
6.6	Evaluation Metrics	75
	6.6.1 Accuracy	75
	6.6.2 Performance	75
	6.6.3 Reliability	76
	6.6.4 User Satisfaction	77
CONCLUSION		78
7.1	Summary of Work	78
7.2	Significance and Impact	78
7.3	Future Direction	79
	7.3.1 Expand Training Data	79
	7.3.2 Deployment and Clinical Validation	79
	7.3.3 Model Interpretability	79
	7.3.4 Leveraging Multimodal Data	79
	7.3.5 Continual Learning	80
REFERENCES		81
APPENDICES		86