



FINAL YEAR PROJECT REPORT

**DESIGN AND IMPLEMENTATION OF DEEP
LEARNING METHOD FOR IMAGE MINING
FOR FACE DATABASE**

**In fulfillment of the requirement
For degree of
BS (Information Technology)**

By

**MUHAMMAD USMAN
AIMEN MIRZA**

**45909 BSIT
45894 BSIT**

SUPERVISED

BY

DR. HUMERA FAROOQ

**BAHRIA UNIVERSITY (KARACHI CAMPUS)
2016-2020**

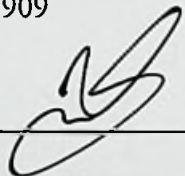
DECLARATION

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.

Name : Muhammad Usman

Reg No. : 45909

Signature :



Name : Aimen Mirza

Reg No. : 45892

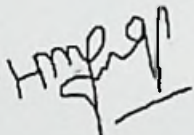
Date :

Aimen

APPROVAL FOR SUBMISSION

We certify that this project report entitled **“DESIGN AND IMPLEMENTATION OF DEEP LEARNING METHOD FOR IMAGE MINING FOR FACE DATABASE”** was prepared by **MUHAMMAD USMAN** and **AIMEN MIRZA** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Information Technology at Bahria University.

Approved by,



Signature : _____

Supervisor: Dr. Humera Farooq

Date : 15 June, 2020

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.

© 2019 Bahria University. All right reserved.

ACKNOWLEDGEMENTS

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express our gratitude to our research supervisor, Dr. Humera Farooq 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 had helped and given us encouragement.

DESIGN AND IMPLEMENTATION OF DEEP LEARNING METHOD FOR IMAGE MINING FOR FACE DATABASE

ABSTRACT

We present a project that proposes an algorithm which is able to segregate random pictures into organized groups. We use the technique of unsupervised learning i.e. clustering for doing so. Our effort makes it possible to cluster all images with respect to the categories we create. Here we will focus on one of the categories i.e. poses.

We are using five datasets named Georgia Tech face database, Yale Face database, CAS-PEAL, MIT-CBCL and Kohn Kanade. However, in this particular piece of work, we will use Georgia Tech face database. We are focussing on mining the unlabelled images having various poses into separate groups. For this we are applying the unsupervised learning technique of machine learning. We are making clusters that will gather similar poses of the subjects. Hierarchical Clustering method is being used for this purpose. The images are inserted into the algorithm; they are then pre-processed so that they all align over a specific set of resolution and image type. After pre-processing the algorithm form clusters and sends images to their respective clusters.

TABLE OF CONTENTS

DECLARATION	ii
APPROVAL FOR SUBMISSION	iii
ACKNOWLEDGEMENTS	vi
ABSTRACT	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xii
LIST OF FIGURES	xiii

CHAPTER

1	INTRODUCTION	15
	1.1 Background	15
	1.2 Problem Statements	15
	1.3 Aims and Objectives	16
	1.4 Scope of Project	16
2	LITERATURE REVIEW	17
	2.1 CAS-PEAL Face Database	17
	2.2 Cohn Kanade Extended	18
	2.3 Georgia Tech Face Database	19
	2.4 Yale Face Database	19
	2.5 MIT-CBCL	20
	2.6 Machine Learning Technique	20

3	DESIGN AND METHODOLOGY	21
3.1	Unsupervised Learning	21
3.2	Clustering	22
3.3	Basic Architecture	25
3.4	Datasets	26
3.4.1	Georgia Tech Face Database	26
3.4.2	CAS-PEAL	26
3.4.3	Yale Face Database	26
3.4.4	MIT-CBCL	26
3.4.5	Kohn Kanade	26
3.5	Features/Categories	27
3.5.1	Poses	27
3.5.2	Facial expressions/Emotions	28
3.5.3	Lighting Condition	28
4	IMPLEMENTATION	29
4.1	Image Mining for Face Database code	29
4.1.1	Load Data	29
4.1.2	Shuffle the Data	30
4.1.3	Reducing Dimensions to StandardScaler.	31
4.1.4	Applying Agglomerative Clustering	31
4.1.5	Loading Image within same dataset to test result	32
4.1.6	Suppling loaded image to predict	32
5	RESULTS	33
5.1	Datasets	33
5.1.1	Sample images in CAS PEAL	35
5.1.2	Poses in CAS PEAL	36
5.1.3	Expression Categories in Cohn Kanade	36
5.1.4	Poses variations in MIT CBCL Database	39
5.2	Features/Categories	39
5.3	Outcomes	42
5.4	Limitations	42

6	CONCLUSION	43
	6.1 Conclusion	43
	6.2 Future work	43
	REFERENCES	44