**Minor Project Report**

**on**

**Deep Learning for Emotion Recognition in Cartoons**

Submitted to Guru Gobind Singh Indraprastha University, Delhi (India)

In partial fulfillment of the requirement for the award of the degree

of

**Bachelor of Technology**

**in**

**Information Technology**

Under the guidance of

**Ms. Preeti Sehrawat**

**Assistant Professor**

**Information Technology**

Submitted By:

**Shubham Kumar Abhishek Sharma Arjun Sharma (00815007721) (00115007721) (00415007721)**

**MAHARAJA SURAJMAL INSTITUTE OF TECHNOLOGY**

**C4, JANAKPURI, NEW DELHI – 110058**

(GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY,

DWARKA, NEW DELHI)

**(December 2023)**

**CANDIDATE’S DECLARATION**



It is hereby certified that the work which is being presented in the B.Tech. Minor Project report entitled **Deep Learning for Emotion Recognition in Cartoons** in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** and submitted in the **Department of Information Technology, New Delhi (Affiliated to Guru Gobind Singh Indraprastha University, New Delhi)** is an authentic record of our work carried out during a period from **August 2023 to December 2023** under the guidance of **Ms. Preeti Sehrawat (Assistant Professor).**

The matter presented in the B.Tech. Minor Project Report has not been submitted by us for the award of any other degree of this or any other institute.

**Shubham Kumar Abhishek Sharma Arjun Sharma (00815007721) (00115007721) (00415007721)**

This is to certify that the above statement made by the candidates are correct to the best of my knowledge. They are permitted to appear in the External Minor Project Examination.

**Ms. Preeti Sehrawat Prof. Prabhjot Kaur (Assistant Professor) (HOD, IT Department)**

The B.Tech. Minor Project Viva-Voce Examination of **Shubham Kumar (00815007721), Abhishek Sharma (00115007721), Arjun Sharma (00415007721)** has been held on **19.12.2023.**

**Project Coordinator (Signature of External Examiner)**

**ACKNOWLEDGEMENT**



We express our deep gratitude to **Ms. Preeti Sehrawat**, (Assistant Professor), Department of Information Technology for her valuable guidance and suggestions throughout our project work. We are thankful to **Dr. Preeti Rathee**, Project Coordinator for her valuable guidance.

We would like to extend our sincere thanks to the **Head of Department, Prof. Prabhjot Kaur** for her time-to-time suggestions to complete our project work. We are also thankful to **Prof. Archana Balyan, Director (O) of MSIT** for providing us with the facilities to carry out our project work.

**Shubham Kumar Abhishek Sharma Arjun Sharma (00815007721) (00115007721) (00415007721)**

**ABSTRACT**



Emotion Recognition is a field that computers are getting very good at identifying; whether it’s through images, video or audio. Emotion Recognition has shown promising improvements when combined with classifiers and Deep Neural Networks showing a validation rate as high as 59% and a recognition rate of 56%. The focus of this dissertation will be on facial based emotion recognition. This consists of detecting facial expressions in images and videos. While the majority of research uses human faces in an attempt to recognize basic emotions, there has been little research on whether the same deep learning techniques can be applied to faces in cartoons. The system implemented in this paper, aims to classify at most three emotions (happiness, anger and surprise) of the 6 basic emotions proposed by psychologists Ekman and Friesen, with an accuracy of 80% for the 3 emotions. Showing promise of applications of deep learning and cartoons. This project is an attempt to examine if emotions in cartoons can be detected in the same way that human faces can.

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Fig. No.** | **Title of the Figure** | **Page Number** |
| 1.1 | [An example of the McCulloch-Pitts (MCP) neuron](#_1t3h5sf) | 2 |
| 1.2 | [An example of the Perceptron](#_4d34og8) | 3 |
| 1.3 | [The XOR problem](#_2s8eyo1) | 4 |
| 1.4 | [A Multi-Layer Perceptron](#_17dp8vu) | 5 |
| 1.5 | [Convolutional Neural Network](#_3rdcrjn) | 6 |
| 1.6 | [Recurrent Neural Network & Vanishing Gradient Problem](#_26in1rg) | 7 |
| 1.7 | [LSTM unit](#_lnxbz9) | 7 |
| 2.1 | An illustrated example of parameter sharing | 17 |
| 2.2 | [An example of a 2D convolution](#_35nkun2) | 19 |
| 2.3 | Max pooling | 20 |
| 2.4 | [Fully Connected Network & ReLU.](#_1ksv4uv) | 21 |
| 2.5 | [An illustrated example of the softmax function](#_44sinio) | 22 |
| 2.6 | [AlexNet architecture](#_2jxsxqh) | 24 |
| 2.7 | [ImageNet classification on Cartoons](#_z337ya) | 26 |
| 3.1 | [*Trello* Board](#_3j2qqm3) | 29 |
| 3.2 | [This project’s *Trello* board](#_1y810tw) | 29 |
| 4.1 | [A scene from *Tom & Jerry*](#_4i7ojhp) | 33 |
| 4.2 | [*YouTube* results for the query ‘Tom & Jerry’](#_2xcytpi) | 34 |
| 4.3 | Sample set of videos from the Joni Valentayn YouTube channel | 35 |
| 4.4 | [Different Haar-like features](#_1ci93xb) | 36 |
| 4.5 | [Haar-like features detecting features in a face](#_3whwml4) | 36 |
| 4.6 | [The architecture of the 1st step process, dataset construction & segmentation](#_2bn6wsx) | 38 |
| 4.7 | [The architecture of the 2nd step process.](#_qsh70q) | 39 |
| 4.8 | [Haar cascade training for positive images](#_3as4poj) | 42 |
| 4.9 | [Before and after segmenting a region of a face](#_1pxezwc) | 44 |
| 4.10 | [Result after multi-class classification](#_49x2ik5) | 46 |
| 4.11 | [Convolution visualizations](#_2p2csry) | 47 |
| 5.1 | Results Graphs, Run 1 | 53 |
| 5.2 | Results Graphs, Run 2 | 55 |
| 5.3 | Results Graphs, Run 3 | 56 |
| 5.4 | Results Graphs, Run 4 | 58 |
| 5.5 | Results Graphs, Run 5 | 59 |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Title of the Table** | **Page Number** |
| 2.1 | [ImageNet classification of the zebra by (Randomlists.com,](#_147n2zr) [n](#_3o7alnk)[.d) in Figure](#_147n2zr) [2.](#_2jxsxqh)[6.](#_147n2zr) . | 24 |
| 5.1 | [Hyperparameter Table, Run 1.](#_23ckvvd) | 52 |
| 5.2 | [Results Table, Run 1.](#_ihv636) | 53 |
| 5.3 | [Hyperparameter Table, Run 2.](#_32hioqz) | 54 |
| 5.4 | [Results Table, Run 2.](#_1hmsyys) | 54 |
| 5.5 | [Hyperparameter Table, Run 3.](#_41mghml) | 55 |
| 5.7 | [Results Table, Run 3.](#_2grqrue) | 56 |
| 5.7 | [Hyperparameter Table, Run 4.](#_vx1227) | 57 |
| 5.8 | [Results Table, Run 4.](#_3fwokq0) | 57 |
| 5.9 | [Hyperparameter Table, Run 5.](#_1v1yuxt) | 58 |
| 5.10 | [Results Table, Run 5.](#_4f1mdlm) | 59 |
| 5.11 | [Algorithms with the best loss accuracy out of all 5 runs.](#_2u6wntf) | 60 |
| 6.1 | [The original risk matrix from project proposal](#_19c6y18) | 62 |
| B.1 | [The first 32 *Tom & Jerry* episodes of the images that are included in the dataset of](#_3tbugp1) [this project.](#_3tbugp1) | 65 |
| B.2 | [The last 32 *Tom & Jerry* episodes of the images that are included in the dataset of](#_28h4qwu)  [this project.](#_28h4qwu) | 66 |

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| **Abbreviation** | **Full form of Abbreviation** |
| **AFEW** | **A**cted **F**acial **E**xpression in the **W**ild |
| **ANN** | **A**rtificial **N**eural **N**etwork |
| **CNN** | **C**onvolutional **N**eural **N**etwork |
| **DBN** | **D**eep **B**elief **N**etwork |
| **FER-2013** | **F**acial **E**xpression **R**ecognition-**2013** |
| **FFNN** | **F**eed **F**orward **N**eural **N**etwork |
| **HDF5** | **H**ierarchical **D**ata **F**ormat **5** |
| **HCI** | **H**uman **C**omputer **I**nterface |
| **ILSVRC** | **I**mageNet **L**arge **S**cale **V**isual **R**ecognition **C**hallenge |
| **IRNN** | **I**dentity **R**ecurrent **N**eural **N**etwork |
| **MCP** | **M**c**C**ulloch–**P**itts Neuron |
| **MGM** | **M**etro-**G**oldwyn-**M**ayer |
| **MLP** | **M**ulti–**L**ayered **P**erceptrons |
| **NLP** | **N**atural **L**anguage **P**rocessing |
| **NTM** | **N**eural **T**uring **M**achine |
| **NAG** | **N**esterov **A**ccelerated **G**radient |
| **OpenCV** | **O**pen **C**omputer **V**ision library |
| **PSF** | **P**ython **S**oftware **F**oundation |
| **LSTM** | **L**ong **S**hort **T**erm **M**emory |
| **ReLU** | **Re**ctified **L**inear **U**nit |
| **RNN** | **R**ecurrent **N**eural **N**etwork |
| **RMS** | **R**oot **M**ean **S**quare |
| **SFEW** | **S**tatic **F**acial **E**xpression in the **W**ild |
| **SDLC** | **S**oftware **D**evelopment **L**ife **C**ycle |
| **SGD** | **S**tochastic **G**radient **D**escent |

**CONTENTS**



**Candidate Declaration II**

**Acknowledgement III**

**Abstract IV**

**List of Figures V**

**List of Tables VI**

**List of Abbreviations VII**

**Table of Content VIII**

**1 Introduction 1**

1.1 Outline1

1.2 History of Deep Learning1

1.3 History of Emotion Recognition8

1.4 History of Animated Cartoons10

1.5 Aim11

1.6 Objectives11

1.7 Structure of the rest of the report12

**2 Background 14**

2.1 Related Work14

2.1.1 Emotion Recognition14

2.1.2 Animated Cartoons15

2.2 Convolutional Neural Networks16

2.3 ImageNet23

2.4 Recurrent Neural Networks25

**3 Methodology 28**

3.1 Project Management28

3.1.1 Trello28

3.2 Software Development30

3.3 Research Methods30

**4 Implementation 32**

4.1 Requirements32

4.1.1 Choice of Deep Neural Network32

4.1.2 Choice of Animated Cartoon32

4.1.3 Tom & Jerry32

4.1.4 Dataset Gathering33

4.1.5 Face Segmentation35

4.1.6 Haar-like features35

4.2 Design37

4.2.1 Choice of emotions37

4.2.2 Design of the artifact37

4.2.3 Design of the Convolutional Neural Network39

4.3 Development40

4.3.1 Tools40

4.3.1.1 Python40

4.3.1.2 OpenCV40

4.3.1.3 Keras42

4.3.2 Cartoon Face Segmentation43

4.3.3 Tom & Jerry Image Dataset44

4.3.4 Training, Classification and Visualisation44

**5 Testing and Evaluation 48**

5.1 Preparation48

5.2 Optimisation Algorithms49

5.2.1 Stochastic Gradient Descent49

5.2.2 Adagrad50

5.2.3 Adadelta50

5.2.4 RMSProp51

5.2.5 Adam51

5.3 Results52

5.3.1 Run 152

5.3.2 Run 254

5.3.3 Run 355

5.3.4 Run 457

5.3.5 Run 558

**6 Reflection 61**

**A Cartoon Face Segmentation 63**

**B List of Episodes in Dataset 65**

**C Classification 67**

**D** **Emotion Classification Results 69**

**References 71**