VISVESVARAYA TECHNOLOGICAL UNIVERSITY

JNANA SANGAMA, BELAGAVI– 590018, KARNATAKA, INDIA

A PROJECT REPORT
ON

"DETECTION OF GENDER, AGE AND EMOTION OF A HUMAN IMAGE USING FACIAL FEATURES"

Submitted in partial fulfillment of the requirements for the award of
BACHELOR OF ENGINEERING
IN
COMPUTER SCIENCE & ENGINEERING

Submitted By

Name	USN
VINUTHA RAI K	4VP17CS087
ANVITHA P	4VP18CS400
DEEPIKA B	4VP18CS401
SHAMITHA Y	4VP18CS403

Under the Guidance of

Mrs. Savitha M
Assistant Professor
Department of Computer Science & Engineering
Vivekananda College of Engineering & Technology

DEPARTMENT OF COMPUTER SCIENCE &ENGINEERING
VIVEKANANDA COLLEGE OF ENGINEERING & TECHNOLOGY
[A Unit of Vivekananda Vidyavardhaka Sangha, Puttur (R)]
Affiliated to Visvesvaraya Technological University and Approved by AICTE New Delhi & Govt. of Karnataka
Nehru Nagara, Puttur – 574 203, DK, Karnataka, India

June 2021

VIVEKANANDA COLLEGE OF ENGINEERING & TECHNOLOGY

[A Unit of Vivekananda Vidyavardhaka Sangha, Puttur (R)]
Affiliated to Visvesvaraya Technological University and Approved by AICTE New Delhi & Govt. of Karnataka
Nehru Nagara, Puttur – 574 203, DK, Karnataka, India

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

Certified that the project work entitled "**Detection of Gender, Age and Emotion of a Human Image Using Facial Features** " is carried out by **Miss. Vinutha Rai K, Miss. Anvitha P, Miss. Deepika B, and Miss. Shamitha Y** bearing USNs **4VP17CS087, 4VP18CS400, 4VP18CS401 and 4VP18CS403** respectively, bonafide students of **Vivekananda College of Engineering & Technology, Puttur**, in partial fulfillment for the award of **Bachelor of Engineering** in **Computer Science & Engineering** of the **Visvesvaraya Technological University, Belagavi** during the year 2020 – 21. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library.

The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the said Degree.

Signature of the Guide **Mrs. Savitha M**	**Signature of the HOD** **Prof. Krishna Mohana A J** HEAD OF THE DEPARTMENT Computer Science & Engineering VCET, Puttur, D. K. - 574203	**Signature of the Principal** **Dr. Mahesh Prasanna K** PRINCIPAL Vivekananda College of Engg. & Technology, PUTTUR 574 203.

EXTERNAL VIVA

Name of the Examiners

Signature with date

1..... .....

2..... .....

ABSTRACT

The main motive is to develop an automatic age and gender estimation method towards human faces which will continue to possess an important role in computer vision and pattern recognition. Apart from age estimation, facial emotion recognition also plays an important role in computer vision. Non-verbal communication methods such as facial expressions, eye movement and gestures are used in many applications of human computer interaction. In order to create computer modeling of humans age, gender and emotions a plenty of research has been accomplished. But it is still far behind the human vision system. In this project, we propose a Convolutional Neural Network (CNN) based architecture for age & gender classification. The architecture is trained to label the input images into 8 labels of age and 2 labels of gender. In order for computer modeling of human's emotions we are planning to predict human emotions using deep CNN and observe how emotional intensity changes on a face from low level to high level of emotion.

ACKNOWLEDGEMENT

We take this opportunity to express our deep heartfelt gratitude to all those people who have helped us in the successful completion of the project.

First and foremost, we would like to express our sincere gratitude towards our project guide **Mrs. Savitha M.** for providing excellent guidance, encouragement and inspiration throughout the project work. Without her invaluable guidance, this work would never have been a successful one.

We would like to express our sincere gratitude to Head of the Department of Computer Science & Engineering, **Prof. Krishna Mohana A J** for his guidance and inspiration.

We would like to thank the Principal, **Dr. Mahesh Prasanna K** for providing all the facilities and a proper environment to work in the college campus.

We are thankful to all the teaching and non-teaching staff members of Computer Science & Engineering Department for their help and needed support rendered throughout the project.

DECLARATION

We, **Vinutha Rai K (4VP17CS087), Anvitha P (4VP18CS400), Deepika B (4VP18CS401), and Shamitha Y (4VP18CS403)** students of eight semester B. E. in Computer Science & Engineering, **Vivekananda College of Engineering & Technology**, Puttur, hereby declare that the project work entitled **"Detection of Gender, Age and Emotion of a Human Image Using Facial Features "** has been carried out and duly executed by us at VCET, Puttur, under the guidance of **Mrs. Savitha M**, Assistant Professor, Department of Computer Science & Engineering, Vivekananda College of Engineering & Technology, Puttur, and submitted in partial fulfillment of the requirements for the award of degree in **Bachelor of Engineering in Computer Science &Engineering** by **Visvesvaraya Technological University**, Belagavi during the academic year 2020-21.

Names of the Students	USN	Signature with Date
VINUTHA RAI K	4VP17CS087	
ANVITHA P	4VP18CS400	
DEEPIKA B	4VP18CS401	
SHAMITHA Y	4VP18CS403	

Date:

Place: VCET, Puttur

TABLE OF CONTENTS

Title	Page No.
List of figures	**II**
List of tables	**III**
CHAPTER 1 INTRODUTION	**01**
1.1 Introduction to the project	02
1.2 Literature Survey	02
1.3 Comparison with existing systems	03
1.4 Proposed System	04
CHAPTER 2 REQUIREMENT SPECIFICATION AND ANALYSIS	**05**
2.1 Introduction	06
2.2 Functional requirements	06
2.2.1 Module to train the CNN model	06
2.2.2 Web interface for input image to be tested	07
2.2.3 Interface to display the output	07
2.3 Non-functional requirements	07
2.3.1 Stability	07
2.3.2 Reliability	07
2.3.3 Availability	07
2.3.4 Usability	08
2.3.5 Performance	08
2.3.6 Security	08
2.3.7 User-friendly	08
2.3.8 User interface requirements	08
2.3.9 Result Display	08
2.4 Hardware requirements	09
2.5 Software requirements	09
2.5.1 Windows 10	09
2.5.2 Jupiter notebook	09
2.5.3 Python	09
2.5.4 Google collab	10
CHAPTER 3 SYSTEM DESIGN	**11**
3.1 Introduction	12
3.2 Overview of the proposed system	12
3.3 Architecture Design	12
3.3 Dataflow Diagram	14
3.4 Sequence Diagram	16

3.6 Flow chart 18

3.7 Use case diagram 19

CHAPTER 4 SYSTEM IMPLEMENTATION 21

4.1 Introduction 22

4.2 Algorithm Used 22

4.2.1 CNN model creation 22

4.2.2 Xception Architecture 23

4.2.3 VGG 16 Architecture 24

4.3 Implementation requirements 25

4.4 Work flow 25

4.5 User Interface 26

CHAPTER 5 SYSTEM TESTING AND 30 EXPERIMENTAL RESULTS27

5.1 Introduction 28

5.2 Types of Tests 28

5.2.1 Unit testing 28

5.2.2 Integration testing 29

5.2.3 System test 29

5.2.4 White box testing & black box testing 29

5.2.5 Acceptance testing 29

5.3 Test Result 30

5.4 Experimental Results 30

5.4.1 Datasets 30

CHAPTER 6 SCREENSHOTS 33

CHAPTER 7 CONCLUSION AND SCOPE FOR FUTURE ENHANCEMENT 36

7.1 Conclusion 37

7.2 Scope for Future Enhancement 37

REFERENCES 38

LIST OF FIGURES

Figure No.	Title	Page No.
3.1	System Architecture Design	12
3.2	Level 0 data flow diagram of the system	15
3.3	Level 1 DFD of the system	15
3.4	Level 2 DFD of the system	16
3.5	Sequence diagram of facial expression recognition system	17
3.6	Sequence diagram of age and gender	17
3.7	Sequence diagram of facial expression recognition	18
3.8	Train CNN-Xception Model	19
3.9	Use-case Diagram	20
4.1	CNN Architecture	23
4.2	Xception Architecture	23
4.3	VGG-16 Architecture	24
4.4	VGG-16 Architecture Map	24
5.1	Sample images from FER-2013 Dataset	31
5.2	Confusion Matrix	31
5.3	Sample images from IMDB-WIKI dataset	32
6.1	Detection of age, gender with expression angry	34
6.2	Detection of age, gender with expression happy	34
6.3	Detection of age, gender with expression neutral	35
6.4	Detection of age, gender with expression happy	35

LIST OF TABLES

Table No.	Title	Page No.
5.1	Test Result	31

INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1 Introduction to the Project

Age estimation from face images plays an important role in human and computer vision which has many applications in for example, forensics or social media. It can determine the prediction of other biometrics of human and facial attributes tasks such as gender, ethnicity, hair color and expressions. Large amount of research has been conducted to determine age estimation using facial features. Different public standard datasets which can be used for real age estimation which permit public performance comparison of the proposed methods. The facial expressions can be recognized using non-verbal communication between humans, along with the interpretation of facial expressions have been widely studied. Facial expression plays an important role in human interaction, Facial Expression Recognition (FER) algorithm with the help of computer vision which helps in applications such as human-computer interaction and data analytics.

1.2 Literature Survey

Vladimir khryashchev, Alexander Ganin, Olga Stepanova, Anton Lebedev [1] proposed a system Age estimation from face image. The real-time audience measurement system consists of five consecutive stages. Face detection, face tracking, gender recognition, age classification and in-cloud data statistics analysis, the challenging part such system is age estimation algorithm on the basis of machine learning methods. The face aging process is determined by different factors: genetic, lifestyle, expression and environment. Adaptive feature extraction based on local binary pattern and support vector machine classification.

Xiaofeng Wang, Azliza Mohd Ali, Plamen Angelov [2] proposed a system Age estimation algorithm. This paper approaches to classify gender and age from images of human faces which is an essential part of our method for autonomous detection of anomalous human behavior. Automatic detection can help to recognize human behavior which later can assist in investigating suspicious events. Transformation to LBP feature space and SVM training procedure are used for binary classifier construction. To predict direct age binary classifier

outputs are statistically analyzed and the most probable age becomes the algorithm output. Data sets with labels of gender and age exist of face images, they build one dataset named GAFace and applied our proposed method to this dataset achieving excellent results and robustness gender classification: 90.33% and age classification: 80.17% accuracy approaching human performance.

Gil Levi, Tal Hasneer [3] proposed a system Age and Gender Classification using Convolutional Neural Networks. One of the first applications of convolutional neural networks (CNN) is perhaps the LeNet-5 network described by for optical character recognition. Compared to modern deep CNN, their network was relatively modest due to the limited computational resources of the time and the algorithmic challenges of training bigger networks. In this paper they used Gaussian Mixture models were used to represent the distribution of fascial patches for age classification. Deep CNN have additionally been successfully applied to applications including human pose estimation, face parsing, facial key point detection, speech recognition and action classification. To our knowledge, this is the first report of their application to the tasks of age and gender classification from unconstrained photos.

Ayesh Gurnani, Kenil Shah, Vandit Gajjar, Viraj Mavani [4] Proposed Classification of Age, Gender, and Emotion. The input image is given to face module in which off-the -shelf face detector Detects the precise face and gives the B-box coordinates. The product of input crop face and reweighted saliency map 30% is fed to Alex Net, which is prepared on ImageNet and finetuned on the datasets of Adience and Affect Net.

D D Pribavkin, P Y Yakimov [5] is proposed Detection of Gender, Age and Emotion of Human Image using Facial Features. The CNN model is built by using VGG-16 architecture. The CNN model is then trained using epochs, where each epoch contains a certain number of training images. To remove distorted and unwanted images, the loss Gauss function is used. For testing, the input image is given by the user. The model makes the predictions to estimate age, gender and emotion of that input image by comparing with the trained images.

1.3 Comparison with Existing Systems

The main aim of the project is to improve the performance of Facial expression, age and gender recognition system. The system is using XCeption architecture which is based on Depth Wise Separable convolutions, that allows us to train much fewer parameters and reduces the training time. We are considering each and every pixels of the region to classify the facial expression so that it will be giving accurate results. In our system we are using early stopping algorithm to

avoid the over fitting.

Xception is a deep Convolutional neural network architecture that involves Depthwise Separable Convolutions. It was developed by Google researchers. The data first goes through the entry flow, then through the middle flow which is repeated eight times, and finally through the exit flow. All Convolution and Separable Convolution layers are followed by batch normalization.

VGG-16 is considered to be one of the excellent vision model architecture till date. Most unique thing about VGG-16 is that instead of having large number of hyper-parameters they focused on having convolutional layer of 3*3 filter with a stride 1 and always used same padding and maxpool layer of 2*2 filter of stride 2. It follows this arrangement of convolution and maxpool layers consistently throughout the whole architecture. In the end it has 2 fully connected layers followed by a softmax for output. The 16 in VGG-16 refers to it has 16 layers that have weights. This network is a pretty large network and it has about 138 million parameters.

In some of the existing facial expression recognition systems, uses DBN which will lead to the poor performance over the large dimensional data and some other existing systems requires more number of parameters to recognize expression correctly. As compared to some of the other existing system, our system is easy to implement and the results are more accurate.

1.4 Proposed System

The objective of the proposed system is to design and develop a detection of Gender, Age and Emotion of a Human Image using Facial Features system for which the datasets are collected from Kaggle website. In this project, a model that learns the features required classifying the fundus images and accurately classifying the majority of proliferative cases and cases with no DR using Convolutional Neural Network is proposed. The major modules of the proposed system are given below:

Train Module: This module initially loads the dataset, and generates a model containing the weights of the neural network.

Test Module: This module is executed only when the user uploads an image dataset file and analyzes the input file. Then, user interface outputs the respective class for the identified image file.

REQUIREMENTS SPECIFICATION

&

ANALYSIS

CHAPTER 2

REQUIREMENTS SPECIFICATION & ANALYSIS

2.1 Introduction

The Requirement Specification and Analysis is a document, which describes completely the external behaviour of the software. It is the first and foremost work of a software developer to study the system to be developed and specify the user requirements before going for the designing part. It is a description of a software system to be developed, laying out functional and non-functional requirement and may include a set of use cases that describe interactions the users will have with the software. System Requirement Specification is the official statement of what is required for the system developer. This chapter includes an exhaustive description about the software requirement specification (SRS) for the Detection of Gender, Age and Emotion using Convolutional Neural Network. SRS includes overall description, functional requirements, supportability, performance requirement, design constraints etc. A functional requirement defines a function of a system and its components. A function is described as a set of inputs, the behaviour, and outputs. Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish. A non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviours.

2.2 Functional Requirements

The complete description of the behaviour of a system developed is explained below, It includes the functional requirements of the system. The functional requirement will describe the features and functionality of the sysem.

2.2.1 Module to train the CNN model

The CNN model is built by using VGG-16 architecture. The CNN model is then trained using epochs, where each epoch contains a certain number of training images. To remove distorted and unwanted images, the loss Gauss function is used. For testing, the input image is given by

the user. The model makes the predictions to estimate age, gender and emotion of that input image by comparing with the trained images.

2.2.2 Web interfaces for input image to be tested

A user interface where the user inputs an image that needs to be tested. User Interface provides a form for the user to accept retinal image file as input and upload an image dataset file. The analysis of input file requires a few seconds and provides output by generating web pages.

2.2.3 Interface to display the output

A CNN model detects Gender, Age and Emotion of human Images and extracts a face region from a webcam video, then classifies the images using CNN model.

2.3 Non-functional requirements

Non-functional requirements are the properties that the product must have. These properties are the characteristics or qualities that make the application attractive, or usable, or fast, or reliable. Non-functional requirements do not alter the product's functionality. Non-functional requirements make up a significant part of the specification.

2.3.1 Stability

Scalability is the capability of a system or process to handle a growing amount of work, or its potential to be enlarged in order to accommodate that growth. In this project, user uploads an image dataset file and analyses the input file. Then, model should be scalable to outputs the respective class for the identified image files.

2.3.2 Reliability

Our proposed system is reliable and can be trusted by the user. Reliability can be defined as the probability that a system will produce correct outputs up to some given time t.

2.3.3 Availability

Availability, in the context of a computer system, refers to the ability of a user to access information or resources in a specified location and in the correct format. An availability plan should clearly provide a strategy for availability control. The system is said to be available to the user when he is accessible for authentication.

2.3.4 Usability

Application is able to perform the task easily and efficiently. This is perhaps the basic requirement of any project model. The interface should be easy to learn how to use and easy to remember how to use. Usability Requirements for an interface design should support the following from the perspective of its primary users:

- **Efficiency of use:** Goals are easy to accomplish quickly and with few or no user errors.
- **Intuitiveness:** The interface is easy to learn and navigate; buttons, headings, and help/error messages are simple to understand.
- **Low perceived workload:** The interface appears easy to use, rather than intimidating, demanding and frustrating.

2.3.5 Performance

Application responds to the user without any delay. User should not wait for a long time to get a response from the System. The model is required to have optimum performance. To train a CNN, we need some amount of time given the dataset size is in terms of gigabytes.

2.3.6 Security

System will use secured database. Normal users can just read the information but they cannot edit or modify anything except their personal and some other information and every user has access constraints.

2.3.7 User friendly

System is very interactive. User should use the system without problems. The goal of a user-friendly product is to provide a good user experience.

2.3.8 User Interface Requirements

The user interface (UI) is the key to system usability. The system includes content presentation, system navigation, and user assistance. A UI requirement defines the rules of engagement for a user interacting with an application or system. In this system user can get notification to the particular scenario.

2.3.9 Result Display

The system proposed in this project, accepts image continuously as the input. After getting the image the facial region will be detected and all the important features will be extracted. The features which are extracted from the image will form a knowledge-base. recognition phase,

captured face image is classified based on the match score from the knowledge-base. The matched label of the Facial Expression, age and gender will be displayed.

2.4 Hardware Requirements

The project uses the following hardware components:

- Intel i3Processor Based Computer or higher.
- Memory:4GB RAM.
- Hard Drive:50GB.
- GPU

2.5 Software Requirements

The project requires the following software to run:

- Operating System: Windows 10
- Jupiter notebook
- Python
- Matplotlib, Tensor flow, pandas, NumPy, Keras2
- Google Collab

2.5.1 Microsoft Windows 10

We are working on operating system Windows 10 for Jupiter notebook. This is an open-source web application that allows editing and running notebook documents via a web browser. Here, the Jupiter notebook coded in python language.

2.5.2 Jupiter notebook

The Jupiter Notebook is an open-source web application that allows creating and sharing documents that contain live code, equations, visualizations and narrative text. Uses include data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and much more.

2.5.3 Python

Python is interpreted, high-level, general-purpose programming language. Python has a design philosophy that emphasizes code readability, notably using significant whitespace. Python is an easy to learn, powerful programming language. It has efficient high-level data structures and a simple but effective approach to object-oriented programming. Python’s elegant syntax and

dynamic typing, together with its interpreted nature, make it an ideal language for scripting and rapid application development in many areas on most platforms. It provides constructs that enable clear programming on both small and large scales. Python feature a dynamic type system and automatic memory management. It supports multiple programming, including object-oriented, imperative, functional and procedural. It also has the comprehensive standard library. Python interpreters are available for many operating systems. Python and CPython are managed by the non-profit Python Software Foundation.

2.5.4 Google Collab

Collab is a free notebook environment that runs entirely in the cloud. Helps to edit documents, the way you work with Google Docs. Collab supports many popular machine learning libraries which can be easily loaded in your notebook.

SYSTEM DESIGN

CHAPTER 3

SYSTEM DESIGN

3.1 Introduction

Facial Analysis has gained much recognition in the computer vision community. Human face contains features that determine age, gender and emotions. This chapter contains information on how the design became and how this was done. System design is a process through which requirements are translated into are presentation of software. Initially, the representation depicts the holistic view of the software. Subsequent refinement leads to design representation that is very close to source code. The importance of the system design can be started with a single word quality.

3.2 Overview of the proposed System

The primary objective of the proposed methodology is to recognize the gender, age and emotion from human face images utilizing the set of facial features in real time application. Feature extraction from face images is an important part of this method. In the proposed system, image file is given as input to the trained CNN- Xception for gender and emotion detection and VGG 16 architecture for age detection. This analyses the input file and detects the gender, age and emotion of a human image using facial features. The output is displayed on the user interface.

3.3 Architecture Design

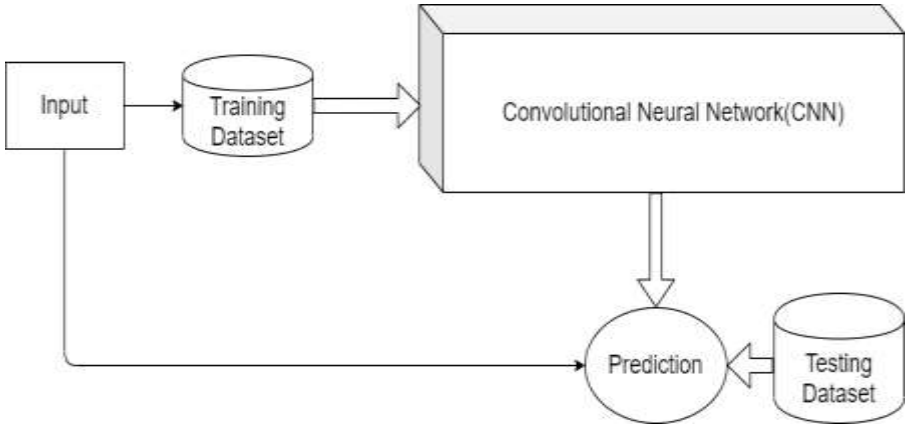

Figure 3.1: System Architecture Diagram

A system architecture is the conceptual model that defines the behaviors, structure, and more views of a system. The input training dataset comprises of the dataset given by the user. This training dataset is given as input to the convolutional neural network. The convolutional network consists of many layers. The input passes through all the layers. Once the training process is done then testing dataset. Finally, the result is predicting the age, gender and emotions of a human image using facial features.

Image: Input image is captured through web camera. Image is obtained from the continuous video input. Web camera is having brilliant resolution and auto colour balance. Image is up to megapixels. Image is in the form of RGB. Proper Resolution and appropriate number of bands are selected.

Face Detection: Face detection is a computer technology being used in a variety of applications that identifies human faces in digital images. Face detection also refers to the psychological process by which humans locate and attend to faces in a visual scene.

The steps in Face Detection are:

- The images are captured continuously using webcam. Then the image is transformed from RGB to Grayscale because it is easy to detect faces in grayscale.
- Next the image manipulation is used in which the resizing, cropping, blurring and sharpening of the image is done if required. Then the image segmentation is done so that the classifier can quickly detect face in the picture.
- The next step is to use Haar features algorithm, which is used for finding face the location of the human faces in a frame or image. The Haar algorithm is used for feature selection for an object in the image, with the help of edge detection, line detection, center detection for detecting eyes, nose, mouth in the picture.
- The next step is to give the coordinates of x, y, w, h which makes a rectangle box in the picture to show the location of the face or to show the region of interest in the image.

Feature extraction: Detecting facial landmarks is a crucial step for face alignment. Facial landmarks are used to localise facial structure such as nose, eyebrows, eyes, mouth and jaw. In machine learning, recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be redundant, then it can be transformed into a reduced set of features (also named a feature

vector). Determining a subset of the initial features is called feature selection. Feature importance scores can be used for feature selection in scikit-learn.

Facial Expression Recognition: Facial expression recognition system is a computer-based technology and therefore, it uses algorithms to instantaneously detect faces, code facial expressions, and recognize emotional states. It does this by analyzing faces in images or video through computer powered cameras embedded in laptops, mobile phones, and digital signage systems, or cameras that are mounted onto computer screens. Facial analysis through computer powered cameras generally follows two steps:

Facial landmark detection: Extracting information about facial features from detected faces. For example, detecting the shape of facial components or describing the texture of the skin in a facial area. Detecting facial landmarks is a crucial step for face alignment. Facial landmarks are used to localise facial structure such as nose, eyebrows, eyes, mouth and jaw.

- **Facial expression classification:** Analyzing the movement of facial features and/or changes in the appearance of facial features and classifying this information into expression-interpretative categories such as facial muscle activations like smile or frown; emotion categories like happiness or anger; attitude categories like (dis)liking or ambivalence.
- **Age and gender classification:** Facial feature such as eyebrows and distance between them, nose, mouth length and face landmarks points are extracted using DLIB library which is present in OpenCV. Then age, gender and emotion features are integrated as one for training.

Output: The detection of age, gender and emotion is based on the voting of CNN. In this step comparison between testing image and trained image is done and also it selects the best possible match of testing image from trained folder and the final results are displayed.

3.4 Data Flow Diagram

A data flow diagram (DFD) is a graphical representation of the flow of data through an information system. DFDs can also be for the visualization of data processing. DFD provides no information about the timing or ordering of processes, or about whether processes will opera in parallel. DFD provides no information about the timing or ordering of processes, or about whether processes will operate in sequence or in parallel. Initially, data is pre-processed and then feature extraction is applied. The output from the feature extraction is then given to the classifier for classification. The input image to be checked for abnormalities is then pre-

processed and result is input to the prediction process which compares the image with the results obtained after classification.

Level 0

Figure 3.2: Level 0 shows the Data Flow diagram of Detection of gender, age, emotion of a human image.

Level 1

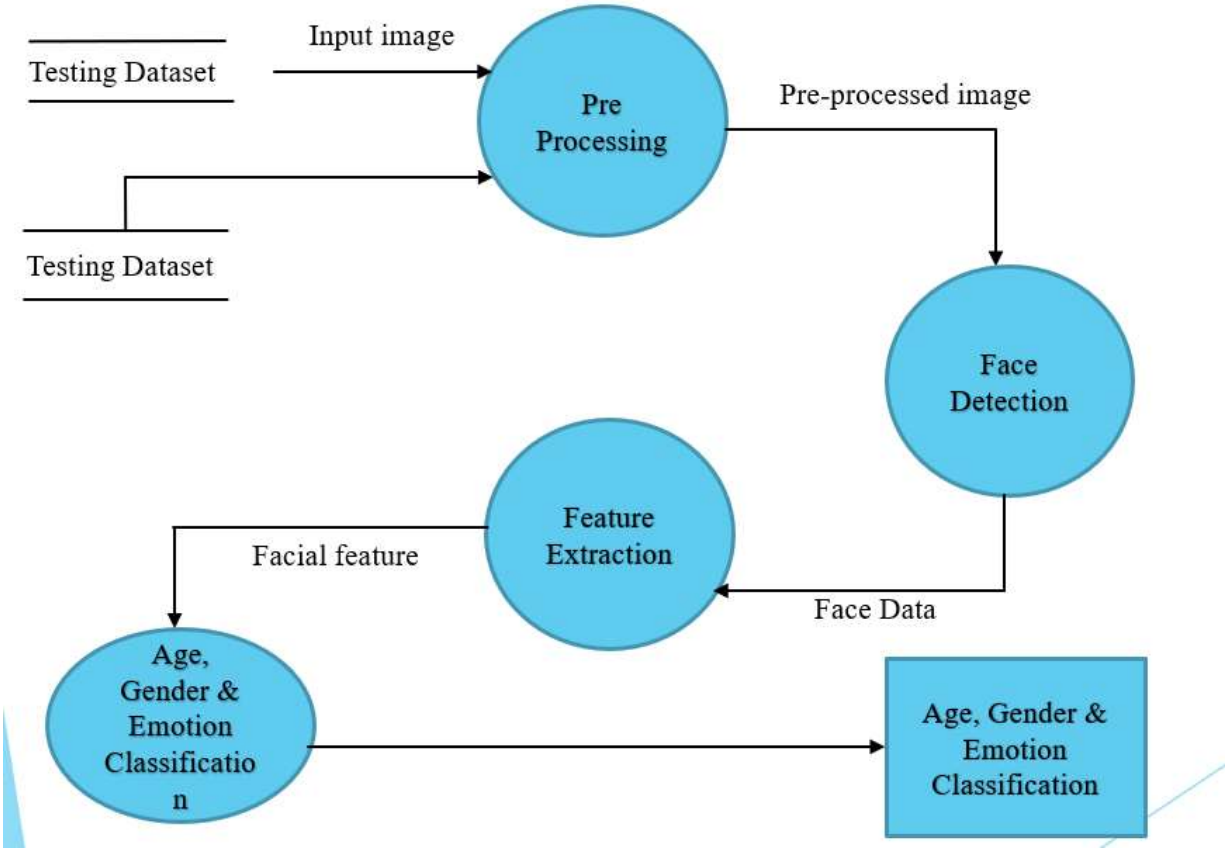

Figure 3.3: Level 1 shows DFD of the system with module.

Level 2

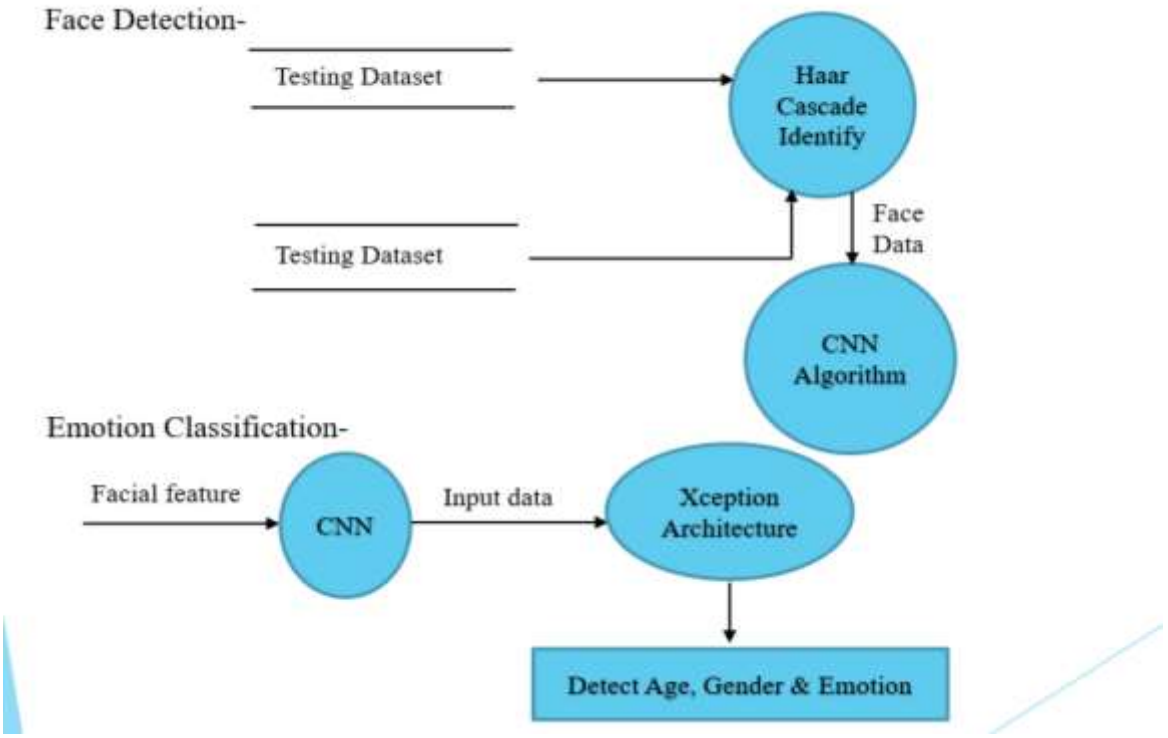

Figure 3.4: Level 2 shows the DFD of the system

3.5 Sequence Diagram

A sequence diagram is an interaction diagram that shows how objects operate with one another and in what order. It is a construction of a message sequence chart. A sequence diagram shows object interactions arranged in time sequence. Sequence diagrams are sometimes called event diagram, event scenarios and timing diagrams.

Each actor as well as system is represented by the vertical line called lifeline and each message is represented by horizontal arrow from sender to receiver. The time will proceed from top to bottom in this diagram. The period of time of an object exists but are not active are shown by dashed line.

The Figure 3.5 shows the facial expression recognition process. When user switches on the camera continuous image is captured from the video input. All captured images which are called trained images are stored in the database. Once feature extraction is done, image whichis to be tested is compared with images stored in the database. As per the result appropriate expression is detected.

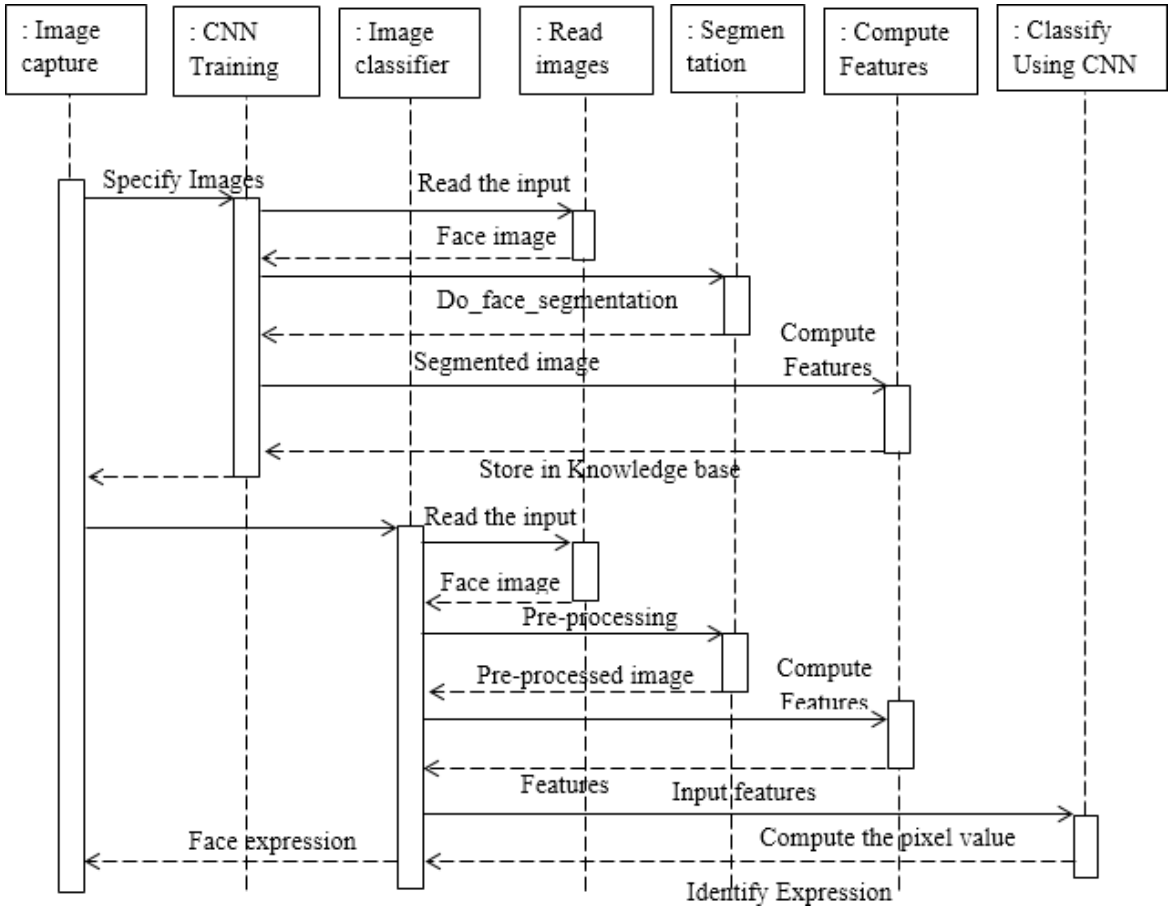

Figure 3.5: Sequence diagram of facial expression recognition system

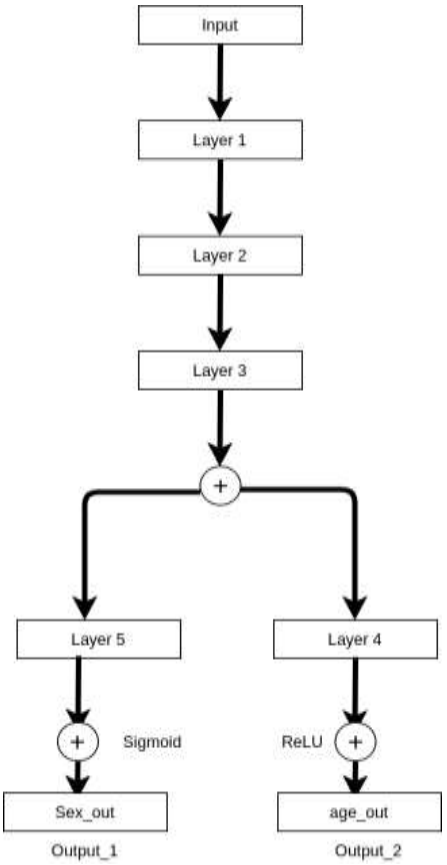

Figure 3.6: Sequence diagram of age and gender detection

3.6 Flowchart

There are 2 phases in building the project, like any neural network projects. In order to detect the gender, age and emotion of human facial features, first we need to build a CNN-Xception and train the same with labeled facial feature image dataset. We employ supervised learning method for the training of the Xception. In our project, we considered 7 well known emotion like angry, disgust, fear, happy, sad, surprise, neutral 2 gender like man and woman. For each stage, we consider about 1000 image files belonging to those stages for the training. The dataset is then used to train the Xception model required for future stage recognition. We developed our project using Tenser flow libraries.

Figure 3.7 shows the training phase of the Convolutional Neural Network. First, we load the dataset for training, i.e., image dataset is loaded for the training of the model, and then we generate the CNN model for each of the training samples.

Figure 3.8 Represents the trained model for stage classification. The trained model uses its training experience on image file and analyses it. Convolution layers of the Xception map the test image file to the corresponding gender, age and emotion of a human image based on facial features. Get image, and classifies image. The gender and emotion of a human images is analyzed by the trained Xception.

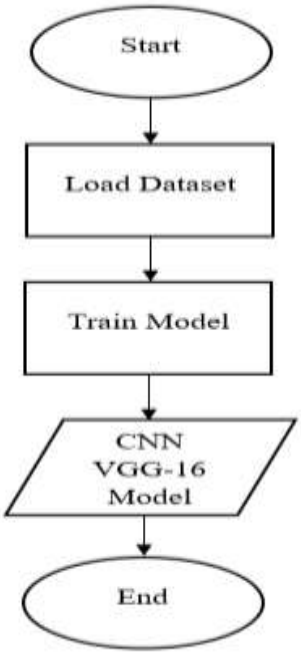

Figure 3.7: Train CNN-Xception Model

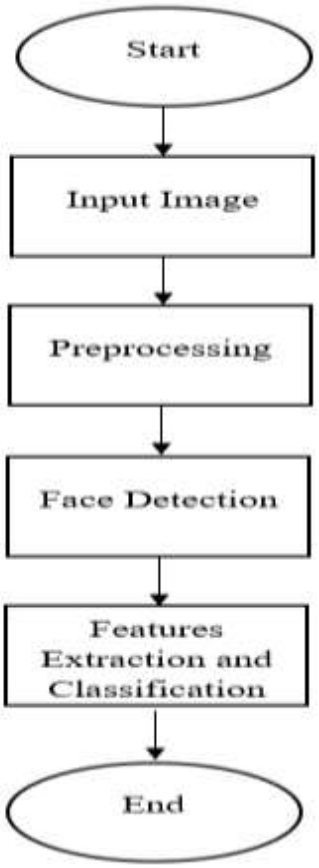

Figure 3.8: Classify image

3.7 Use case Diagram

Use case diagram shows the interaction between user and the system. Use case diagram captures goals of the user and responsibility of system to its users. Use case diagram shows various functions that can be performed by various actors. A use case is an interaction between users and a system. It captures the goal of the users and the responsibility of the system to its users. A user case diagram is a group of actors, a set of use cases enclosed by a system boundary, communication associations between the actors and use cases and generalization among the use cases. images as input. The user uploads the dataset to train the system along with the test data. The user can view the result displayed by the system.

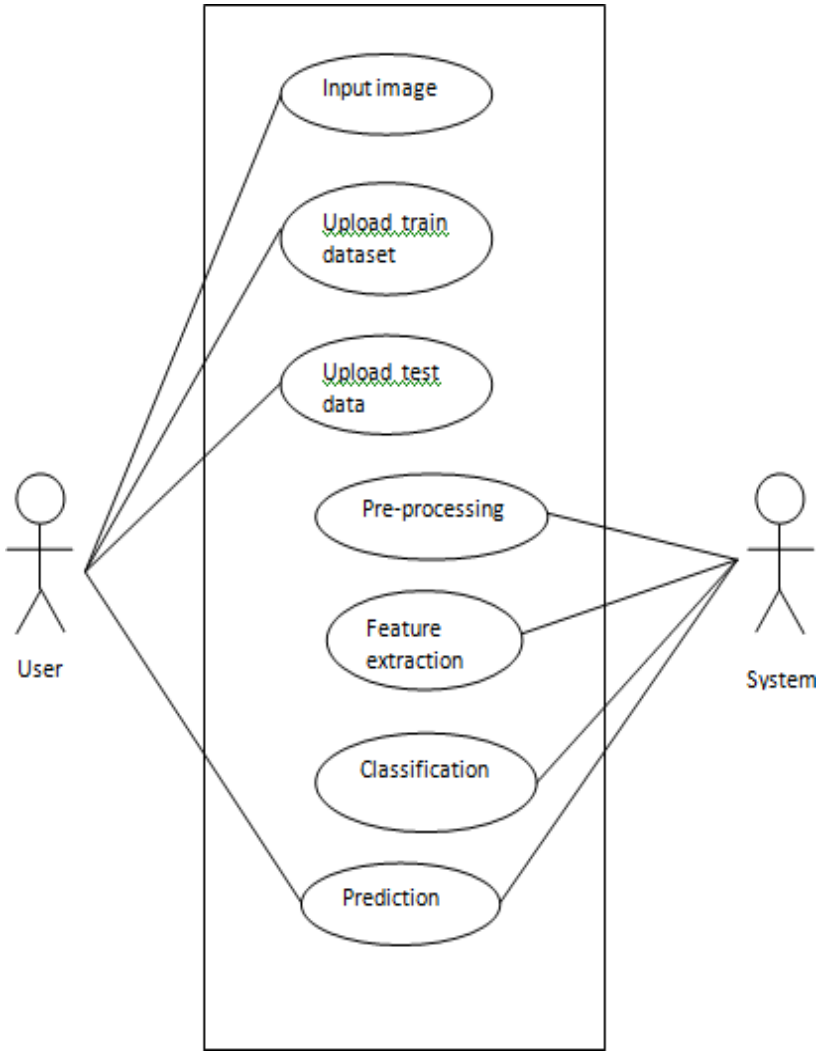

Figure 3.9: Use case diagram

SYSTEM IMPLEMENTATION

CHAPTER 4

SYSTEM IMPLEMENTATION

4.1 Introduction

A crucial phase in the system development life cycle is successful implementation of new system design. Implementations simply mean converting new system design into operation. This is the moment of truth the first question that strikes in every one's mind that whether the system will be able to give all the desires results as expected from system. The term implementation has different meanings, ranging from the conversion of a basic application to a complete replacement of computer system. Implementation is used hereto mean the process of converting a new or revised system design into an operational one.

4.2 Algorithm Used

In our project, we have used Convolution Neural Network model for the genre classification. Convolution Neural Network is a deep learning algorithm which is mainly used to differentiate images based on their spatial features. The calculation runs on Tenser flow stage and depends on the exchange learning philosophy. The classifier has been pre prepared to perform fundamental arrangement dependent on directed preparing technique. This calculation held for the outskirt risk dataset to group different degrees

of dangers. In view of how huge the dataset is and how well it is prepared, the calculation arranges the new pictures. In this stage, we initially introduce the convolutional neural system by characterizing a consecutive constructor. This consecutive constructor contains a direct pile of layers.

4.2.1 CNN Model Creation

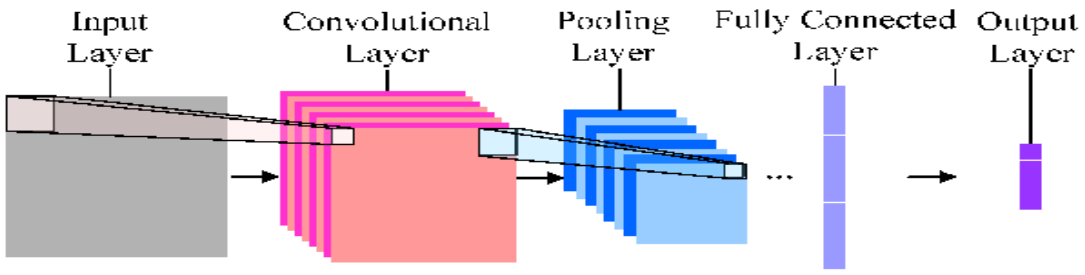

Figure 4.1: Convolution Neural Network Architecture

4.2.2 Xception Architecture

Xception is a deep Convolutional neural network architecture that involves Depthwise Separable Convolutions. It was developed by Google researchers. The data first goes through the entry flow, then through the middle flow which is repeated eight times, and finally through the exit flow. All Convolution and Separable Convolution layers are followed by batch normalization.

Figure 4.2 shows the architecture of Xception model. XCeption is an efficient architecture that relies on two main points:

- Depthwise Separable Convolution.
- Shortcuts between Convolution blocks as in ResNet.

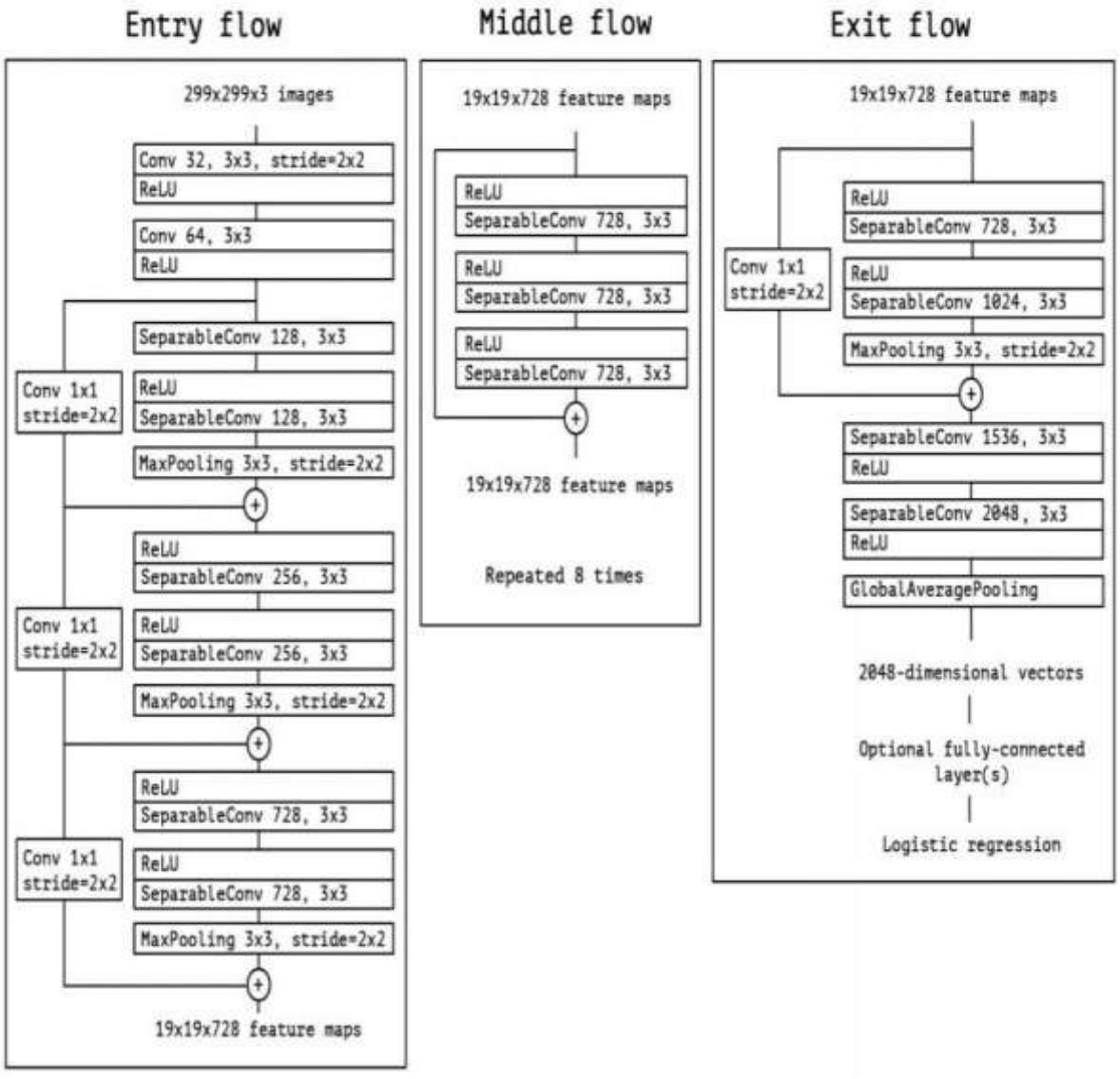

Figure 4.2: Xception Architecture

4.2.3 VGG 16 Architecture

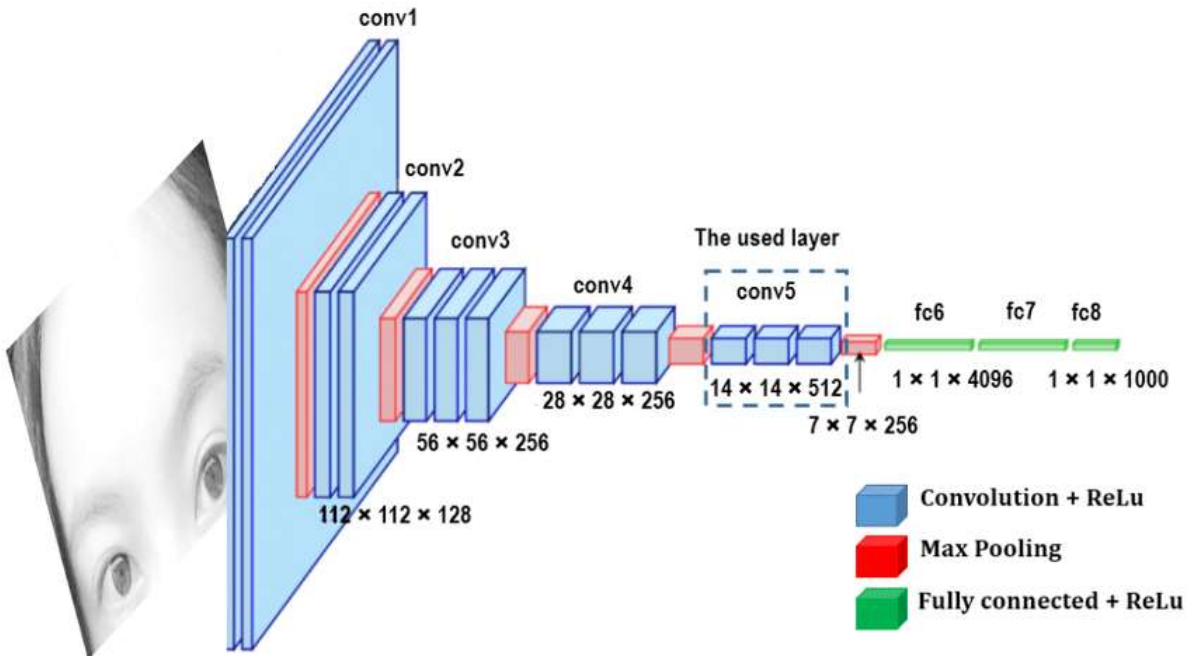

Figure 4.3: VGG 16 Architecture

The input to the network is image of dimensions (224,224,3). The first two layers have 64 channels of 3*3 filter size and same padding. Then after a max pool layer of stride (2,2), two layers which have convolution layers of 256 filter size and filter size (3,4). This followed by a max pooling layer of stride (2,2) which is same as previous layer. Then there are 2 convolution layers of filter size (3,3) and 256 filter. After that there are 2 sets of 3 convolution layer and a max pool layer. Each have 512 filter of (3,3) size with same padding. This image is then passed to the stack of two convolution layer, the filter of size 3*3.

Figure 4.4: VGG 16 Architecture Map

After the stack of convolution and max-pooling layer, we got a (7,7,512) feature map. We flatten this output to make it a (1, 25088) features vector. After this there are 3 fully connected

layer, the first layer takes input from the last feature vector and outputs a (1,4096) vector, second layer also outputs a vector of size (1, 4096) but the third layer output a 1000 channel for 1000 classes of ILSVRC challenges, then after the output of 3rd fully connected layer is passed to SoftMax layer in in order to normalize the classification vector. After the output of classification vector top-5 categories for evaluation. All the hidden layers use ReLU as its activation function. ReLU is more computationally efficient because it results in faster learning and it also decreases the likelihood of vanishing gradient problem.

4.3 Implementation Requirements

XCeption offers an architecture that is made of Depthwise Separable Convolutional blocks + Maxpooling all linked with shortcuts as in ResNet implementations. The specificity of XCeption is that the Depthwise Convolutional is not following by a Pointwise Convolution, but the order is reversed.

First, we imported all the libraries which will need to implement VGG-16. Using Sequential method, we created a sequential model. Sequential model means that all the layers of the model will be arranged in sequence ImageDataGenerator from keras.preprocessing. The objective of ImageDataGenerator is to import data with labels easily into model. It is a very useful class as it has many functions to rescale, rotate, zoom, flip etc. The most useful thing about this class is that it doesn't affect the data stored on the disk. This class alters the data on the go while it to the model.

4.4 Workflow

In the planned design we have the tendency to begin with downloading the image-set from a dataset referred to as IMDB WIKI as a result of it being the largest publicly available dataset with gender and age labels for training. Simultaneously, image-set is downloaded from a dataset referred to as FERC-2013.

4.4.1 Function to encode the labels

After downloading the dataset, we assign labels to this image-set using one-hot encode because CNN does not work with categorical data- variables that contain label values rather than numeric values. To overcome this problem of CNN, we use a one-hot encoding algorithm to convert these label values into numeric values which will be easily processed by CNN.

4.4.2 Resizing and preprocessing the data

Resizing of all the images to fixed pixel 48*48 values are done and preprocessing is done by Data augmentation. Data augmentation is the process of increasing the amount and diversity of data. There will be no collection of new data, rather transforming the already present data.

The most commonly used operations in data augmentation are:

- **Rotation:** Rotation operation as the name suggests, just rotates the image by a certain specified degree.
- **Shearing:** Shearing is also used to transform the orientation of the image.
- **Zooming:** Zooming operation allows us to either zoom in or zoom out.

4.4.3 Extracting and integrating the features

After pre-processing, the facial features such as eyebrows and distance between them, nose, mouth length, and face landmark points are extracted using the DLIB library which is present in OpenCV. Then age, gender and emotion features are integrated as one for training.

4.4.4 Training and Testing

The CNN model is built by using Xception architecture. The CNN model is then trained using epochs, where each epoch contains a certain number of training images. To remove distorted and unwanted images, the loss Gauss function is used. For testing, the input image is given by user. The model makes the predictions to estimate age, gender and emotion of that input image by comparing with the trained images.

4.5 User Interface

We have implemented a web interface for our project which allows the user to interact with the system. The UI provides a form for the user to upload an image dataset file. The analysis of input file requires a few seconds. Once analyzed, UI outputs the respective class for the identified image file. The overall working of the user interface involves following events:

- Input an image file
- Upload the image
- Load model to predict the class
- Display the respective class

SYSTEM TESTING
AND
EXPERIMENTAL RESULTS

CHAPTER 5

SYSTEM TESTING AND EXPERIMENTAL RESULTS

5.1 Introduction

Software testing is the process of checking whether the developed system is working according to the original objectives and requirements. Software testing process commences once the program is created and documentation and related data structures are designed. Software testing is essential for correcting errors. Otherwise, the project is not aid to be complete.

The system should be tested experimentally with test data so as to ensure that the system box according to the required specification. When the system is found working, test it with actual data and check performance. Software testing is a critical element of software quality assurance and represents the ultimate review of specification, design and coding.

The philosophy behind testing is to find errors. Test cases are devised with this purpose in mind. Test case is a set of data that the system will process as normal input.

5.2 Types of Tests

Types of test are: Unit testing, Integration testing, Functional test, System test, Black box testing, Acceptance testing.

5.2.1 Unit Testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application. It is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

5.2.2 Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully. Unit testing, the combination of components are correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components. Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects. The task of the integration test is to check that components or software applications, e.g. components in a software system or one step up software applications at the company level interact without error.

5.2.3 System test

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration-oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

5.2.4 White Box Testing and Black Box testing

White box testing is also known as clear box testing, transparent box testing and glass box testing. White box testing is a software testing approach, which intends to test software with knowledge of internal working of the software. White box testing approach is used in Unit testing which is usually performed by software developers. White box testing intends to execute code and test statements, branches, path, decisions and data flow within the program being tested. Black box testing is a software testing method where testers are not required to know coding or internal structure of the software. Black box testing method relies on testing software with various inputs and validating results against expected output.

5.2.5 Acceptance testing

User knowledge of the inner workings, structure and the language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

5.3 Test Result

Table 5.1: Test Result

Test Case	Description	Input	Expected Result	Obtained Result
1.	Image selection from video source	Video source	Input must be captured from video source	As expected
2.	Image is stored for training dataset	Number of images	Image must be stored in train dataset	As expected
3.	Image is stored for testing dataset	Number of images	Image must be stored in train dataset	As expected
4.	Test image must be compared with test dataset	Load database	System should display the respective age, gender and emotion of human image is called as result.	As expected

5.4 Experimental Results

The system proposed in this project, accepts image continuously as the input. After getting the image the facial region will be detected and all the important features will be extracted. The features which are extracted from the image will form a knowledge-base. recognition phase, captured face image is classified based on the match score from the knowledge-base. The matched label of the Facial Expression, age and gender will be displayed.

5.4.1 Datasets

In this project we used popular FER2013 Kaggle Challenge dataset for emotion detection. The data consists of gray scale face images of 48x48 pixel size. The faces have been automatically registered so that the face is more or less centred and occupies about the same amount of space

in each image. The data set remains quite challenging to use, since there are empty pictures, or wrongly classified images. The Figure 5.1 will be shows some of the images from the dataset.

Figure 5.1: Sample images from fer2013 dataset

In the experimentation, we have used 28709 face images for training and 3589 face images for testing. The recognition accuracy of our system is **64%.** The confusion matrix of this experiment is shown in Figure 5.2.

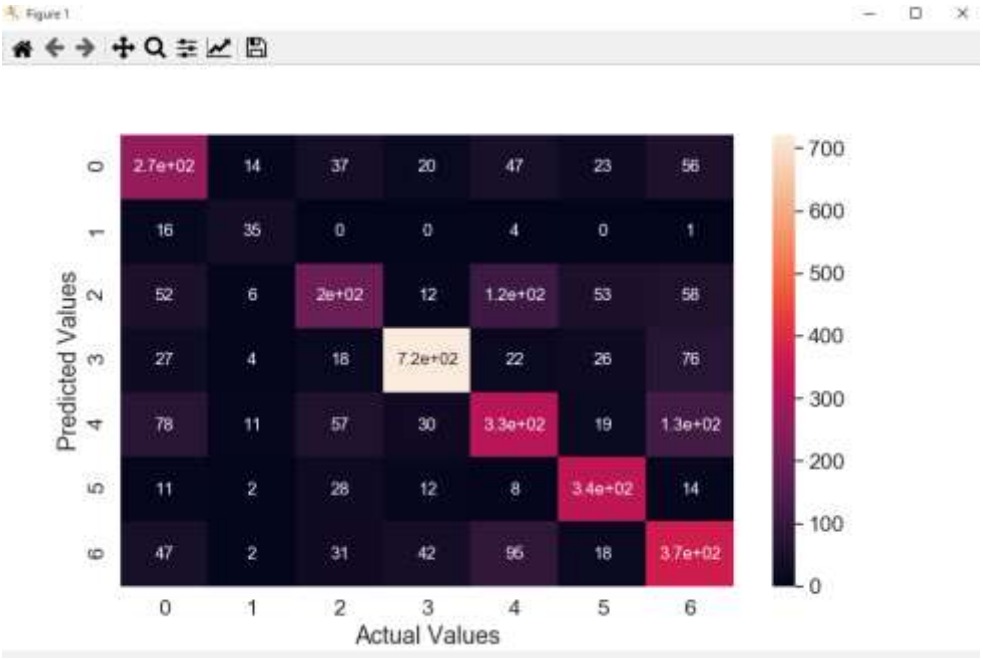

Figure 5.2: Confusion Matrix

In this project to detect age and gender we used IMDB-WIKI dataset from Kaggl**e.** deep learning solution to age estimation from a single face image without the use of facial landmarks and introduce the IMDB-WIKI dataset, the largest public dataset of face images with age and gender labels. If the real age estimation research spans over decades, the study of apparent age estimation or the age as perceived by other humans from a face image is a recent endeavor. We tackle both tasks with our convolutional neural networks (CNNs) of VGG-16 architecture which are pre-trained on ImageNet for image classification. We pose the age estimation problem as a deep classification problem followed by a SoftMax expected value refinement. The key factors of our solution are: deep learned models from large data, robust face alignment, and expected value formulation for age regression. We validate our methods on standard

benchmarks and achieve state-of-the-art results for both real and apparent age estimation.

To the best of our knowledge this is the largest publicly available dataset of face images with gender and age labels for training. We provide pretrained models for both age and gender prediction.

Figure 5.3: Sample images from IMDB-WIKI dataset

SCREENSHOTS

CHAPTER 6

SCREENSHOTS

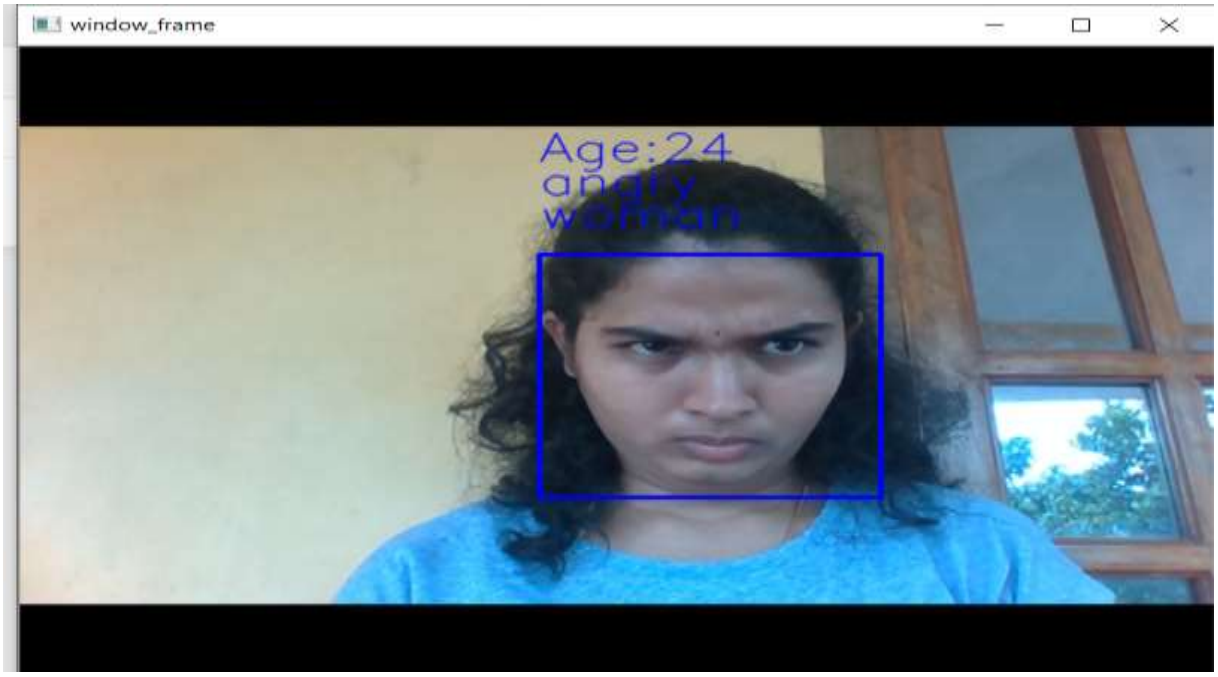

Figure 6.1: Detection of age, gender with expression 'angry'

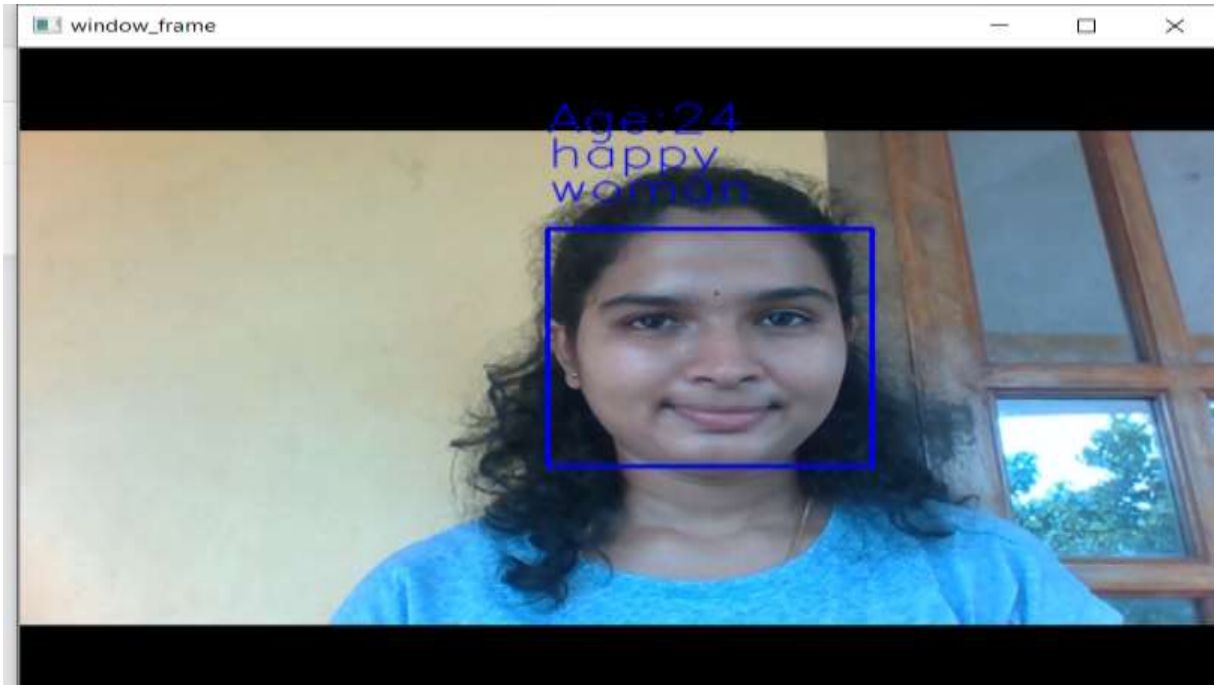

Figure 6.2: Detection of age, gender with expression 'happy'

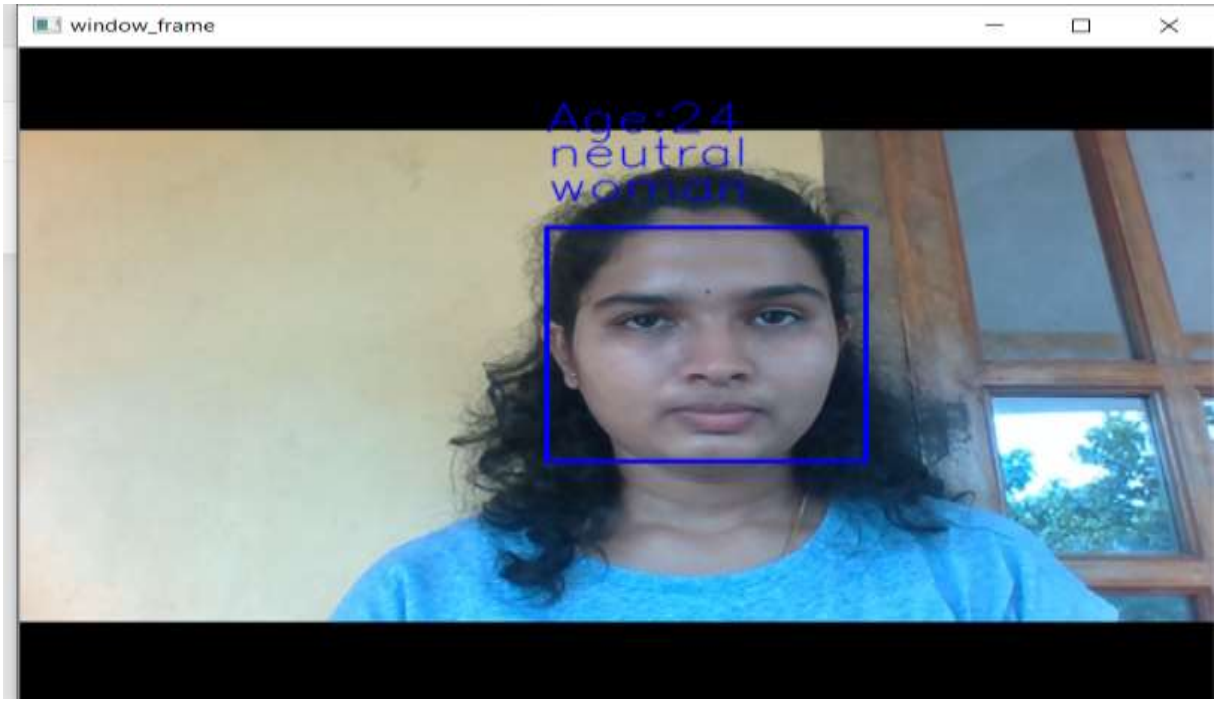

Figure 6.3: Detection of age, gender with expression 'neutral'

Figure 6.4: Detection of age, gender with expression 'happy'

CONCLUSION

AND

SCOPE FOR FUTURE ENHANCEMENT

CHAPTER 7

CONCLUSION AND SCOPE FOR FUTURE ENHANCEMENT

7.1 Conclusion

Here we proposed an accurate and high-speed facial expression recognition system. The major contribution of this system is that it can detect the expression and the results demonstrated so that deep CNN's are capable of learning facial characteristics and improve facial expression recognition. Here the Convolutional networks can intrinsically learn the key facial features by using only raw pixel data. We noticed how the pixels are being activated differently depending on the emotion being labelled. The happiness seems to depend on the pixels linked to the eyes and mouth, whereas the sadness or the anger seems for example to be more related to the eyebrows.

In this System we have used the XGBoost library. This library implements the gradient boosting decision tree algorithm. Boosting is an ensemble technique where new models are added to correct the errors made by existing models. By using this algorithm, we got one of the best model named Xception model. Xception is a deep Convolutional neural network architecture that involves Depth wise Separable Convolutions. In this model the data first goes through the entry flow, then through the middle flow which is repeated eight times, and finally through the exit flow. Xception model allows a shorter training time on GPUs, more images processing per second in real-time prediction, and prevents over fitting. Because of this one we have got the good accuracy and the expressions are recognized accurately.

7.2 Scope for Future Enhancement

Android based implementation can be implemented to get the real time results. The different algorithm can also be used to improve the accuracy. In future accuracy can be increased by expanding the knowledge database by including more face images of gender, different ages and expressions. Disgust and angry expressions are little bit confusing and it effected the accuracy of our system. To find age accurately it is difficult task and it is also effect the accuracy. To avoid this, the model can be trained using the greater number of training data and by considering different parameters.

REFERENCES

[1] Shervin Minaee, Amirali Abdolrashidi, Deep - Emotion: Facial Expression recognition using attentional convolutional network, 2019 3rd IEEE International Conference on Cybernetics (CYBCONF)doi:10.1109/cybconf.2017.7985780

[2] Alen Salihbasi and Tihomir Orehovacki, Development of android application for gender, age and face recognition using opencv, MIPRO 2019, ISSN: 1057-7149, 20 May 2019

[3] Jiu-Cheng Xie, Chi-Man Pun, 'Age and Gender Classification using Convolutional Neural Networks, IEEE Workshop on Analysis and Modelling of Faces & Gestures', IEEE Transactions on Information Forensics and Security, ISSN: 1556-6013,07 March 2019

[4] Gil Levi and Tal Hassner,' Facial Emotion Analysis using Deep Convolutional Neural Network', 2017. International Conference on Signal Processing and Communication (ICSPC), Coimbatore, 2017.

[5] Mostafa Mohammadpour, Seyyed Mohammed. R Hashemi, Hossein Khaliliardali, Mohammad. M Alyan Nezhadi,'Facial Emotion Recognition using Deep Convolutional Networks', 2017 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), 22 December, 2017.

[6] Dehghan, Afshin & G. Ortiz, Enrique & Shu, Guang & Zain Masood, Syed. (2017).DAGER: Deep Age, Gender and Emotion Recognition Using Convolutional Neural Network.

PERSONAL PROFILE

Mrs. Savitha M received her BE Degree in Computer Science & Engineering from KVG college of Engineering Sullia and M. Tech in Computer Science And Engineering from NMAMIT, Nitte. She is a Lecture of department of Computer Science in VCET, Puttur from last 8 years. The subject of Interest includes C++, Java, Data Management, System and Computer Architecture.

Mrs. Savitha M

Project Guide

Vinutha Rai K

4VP17CS087

D/o K Guddappa Rai

Korikaru House, Madavu Post ,Keyyur Village, Puttur,

574210 D.K

E-mail ID: vinuthak0510@gmail.com

Phone No: 9071496110

Anvitha P

4VP18CS400

D/o Shankara Naik P

Periyadka House, Urdur Post, Adoor Village,

Kasaragod – 671543

E-mail ID: anvithap7@gmail.com

Phone No: 9880365302

Deepika B
4VP18CS401
D/o Narayana Bhat G
Banputhadka house
Ukkinadka Post, Enmakaje Village
Kasaragod district, 671552
E-mail ID: deepikabhat1998@gmail.com
Phone No: 7559849534

Shamitha Y
4VP18CS403
D/o Deranna Gowda A
Swasthik Nilaya Yerumbu House
SS Vihar Alike Post
Bantwal TQ, D.K, 574235
Email ID: shamitha.ad@gmail.com
Phone No: 9880993581