Project Phase-2 (19CS703) Report

on

EMOTION RECOGNITION BY INCLUSION OF AGE AND GENDER PARAMETER BY DEEP LEARNING

Submitted to

NMAM INSTITUTE OF TECHNOLOGY, NITTE

(An Autonomous Institution under VTU, Belagavi)

In partial fulfillment of the requirements for the award of the

Degree of Bachelor of Engineering in Computer Science and Engineering

by

ADITHYA HOLLA K	4NM19CS007
ADITYA MURUGAN	4NM19CS010
AKIL RAIF	4NM19CS014
ASHWAMEDH AROTE	4NM19CS031

Under the guidance of

MS. Joylin Priya Pinto

Asst. Professor GD-1

Dept. of CSE, NMAMIT, NITTE



CERTIFICATE

Certified that the project work entitled "Emotion Recognition By Inclusion Of Age And Gender Parameter By Deep Learning" is a bonafide work carried out by Adithya Holla K (4nm19cs007), Aditya Murugan (4nm19cs010), Akil Raif (4nm19cs014), Ashwamed Arote (4nm19cs031) in partial fulfillment for the award of Degree of Bachelor of Engineering in Computer Science and Engineering of the Visvesvaraya Technological University, Belagavi during the year 2022-2023. 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 Phase- 2 (19CS703) prescribed for the said Degree.

Name & Signature of Guide	Signature of HOD	Signature of the Principa
MS. Joylin Priya Pinto		
Asst. Professor GD-1		

External Viva

Name of the Examiners		Signature with Date	
1			
2.			

ACKNOWLEDGEMENT

The satisfaction that accompanies the completion of any task would be incomplete without the mention of all the people, without whom this endeavour would have been a difficult one to achieve. Their constant blessings, encouragement, guidance and suggestions have been a constant source of inspiration

First and foremost, my gratitude to my project guide **Ms Joylin Priya Pinto** for his constant guidance throughout the course of this project Phase-1 and for the valuable suggestions.

I also take this opportunity to express a deep sense of gratitude to the project coordinators for their valuable guidance and support.

I acknowledge the support and valuable inputs given by, **Dr. Jyothi Shetty** the Head of the Department, Computer Science and Engineering, NMAMIT, Nitte.

My sincere thanks to our beloved principal, **Dr. Niranjan N Chiplunkar**for permitting us to carry out this project at our college and providing us with all needed facilities.

Finally, thanks to staff members of the Department of Computer Science and Engineering and our friends for their honest opinions and suggestions throughout the course of our project Phase-1.

ADITHYA HOLLA K (4NM19CS007)
ADITYA MURUGAN (4NM19CS010)
AKIL RAIF (4NM19CS014)
ASHWAMEDH AROTE (4NM19CS031)

ABSTRACT

Automatic gender, age and emotion recognition have relevant to an extension of its usage in various software and hardware, particularly because of the growth of online social networking websites and social media. The advertisements can be specialized based on the age and the gender of the person on the phone. It also can help identify suspects in criminal cases or at least it can minimize the number of suspects. A Convolution Neural Network is a deep neural network (DNN) widely used for the purposes of image recognition and processing and NLP. A convolution neural network architecture is built and the model is trained for gender, age and emotion recognition from images. An android application can be developed implementing the age, gender and emotion recognition. With the phone capturing photos and the frames are pre-processed and fed to the model to accomplish this task. The prediction is displayed accordingly

CONTENT

CHAPTER NO.	CHAPTER NAME	PAGE NO
1	INTRODUCTION	1-4
	 1.1 Introduction 1.2 Overview of Age , Gender , Emotion 1.3 Importance of Age , Gender , Emotion Recognition 1.4 Role Of IT in Age , Gender , Emotion Recognition 1.5 Problem Specification 1.6 Applications Gender, age and emotion recognition 	1 1-2 2 2 3 3
2	LITERATURE SURVEY	4-5
3	PROBLEM DEFINITION	6
4	SYSTEM REQUIREMENTS SPECIFICATION	7
5	SYSTEM DESIGN	8-10
6	IMPLEMENTATION	11-15
7	RESULTS	16-18
8	CONCLUSIONS	19
	REFERENCES	20-21

LIST OF FIGURES

Figure	Description	Page No.
no.		
1	System Architecture	9
2	Working Of Age And Gender Prediction	12
3	Working Of Emotion Prediction	13
4	Working Of Java Script Button	14
5	Ngrok Execution Window	15
6	Flask Output Home page	15
7	Image Browsing Window	16
8	Browsed Image View	16
9	Displaying Predicted Result	17
10	Python Console After Prediction	18
11	Code Snippets	18

INTRODUCTION

1.1 Introduction

In this chapter an overview of gender, age and emotion recognition is provided. The importance of gender, age and emotion recognition and the role of IT in gender, age and emotion recognition is also provided.

1.2 Overview of gender, age and emotion recognition

A change of paradigm is currently underway. Computers have been consistently surpassing humans in many tasks of the day-to-day life in the past decades, with us relying more and more on them. If a bank's database servers would for some reason shutdown and wipe out all of its information, the world economy could collapse. We are in the age of information, and that information is increasingly being collected and stored by machines. Machine Learning uses this information to learn patterns. It can be used to predict the weather, to help on credit approval, or to recommend a visit to the doctor if an irregular heartbeat is detected. But this change of paradigm is currently occurring in a specific sub-field of computer science computer vision, fueled by the advent of Deep Learning and Convolutional Neural Networks. Over the last decade, the rate of image uploads to the internet has been increasing exponentially. Motivated by this and the fact that high computing power is becoming increasingly accessible to the common user, more and more research is being done on computer vision. Problems that were once considered impossible to study due to lack of data or computing power are being attempted now. Cars can drive themselves nowadays, detecting pedestrians and predicting/avoiding accidents. Despite self-driving cars being a hot-topic today, there is perhaps one even hotter – facial analysis. Facial analysis comprehends several tasks, which can be used for several purposes. There are models for the task of face detection, pose estimation, age estimation, face recognition, smile detection, gender recognition, and so on. In many cases, machine estimation has been consistently breaking the human level, achieving outstanding results. There is a sense of urgency in the quest to investigate and develop the best model, fueled by Artificial Intelligence competitions and the tech business world. These systems can be applied to security systems, person identification, human-computer interaction platforms, marketing, etc. We are getting closer, step by step, to become fully connected to technology, which can now lea and estimate our needs and characteristics. Facial analysis plays a key role in this scenario, as a significant part of human communication is nonverbal

1.3 Importance of gender, age and emotion recognition

Facial analysis systems can ultimately change the way we live. Allied to Artificial Intelligence, it can provide personal assistance in every aspect of life. It can estimate our mood upon waking and, connecting it to our habits, age and other factors, provide the perfect meal for the moment, the best music to accompany that meal, and the perfect outfit, linked to the weather prediction. Likewise, upon entering a clothing store, it can automatically suggest the best piece of clothing for you, according to your age and gender, and send to your smartphone a picture of you virtually dressed in such suggestions. Machines will be capable of taking care of the uninteresting part of life, end-to-end. Supported by computer vision and facial analysis, the full connection and understanding of humans will soon happen. In fact, many state-of-the-art models have already outpaced human recognition in many facial analysis tasks.

1.4 Role of IT in gender, age and emotion recognition

The state-of-the-art methods in image-related tasks such as image classification and object detection are all based on Convolutional Neural Networks (CNNs). Human face analysis and soft-biometric classification, has gained more popularity after AlexNet has been introduced by Krizhevsky et al. Such facial soft-biometric include age, gender and facial expression are a topic of interest among many computer vision researchers. Introduction of deep learning to this domain has replaced the need for handcrafted facial attributes and data pre-processing schemes. D-CNN models have been not only successfully applied to Computer Vision

1.5 Problem Specification

Facial analysis systems can ultimately change the way we live. Allied to Artificial Intelligence, it can provide personal assistance in every aspect of life. It can estimate our mood upon waking and, connecting it to our habits, age and other factors, provide the perfect meal for the moment, the best music to accompany that meal, and the perfect outfit, linked to the weather prediction. Machines will be capable of taking care of the uninteresting part of life, end-to-end. Supported by computer vision and facial analysis, the full connection and understanding of humans will soon happen. There are several real-world applications including security and video surveillance, electronic customer relationship management, biometrics, electronic vending machines, humancomputer interaction and entertainment. Deep Learning and Convolutional Neural Networks are a hot-topic in the field of Machine Learning. Many breakthroughs are being reached every other day. Being a part of that is what motivates us to develop this work. Mobile phones are already a huge part of our lives, and combining them with the power of deep learning can create user experiences that delight and impress users. There is no proper android application implementing facial recognition combining gender, age and emotion of a human face using the Convolution Neural Network.

1.6 Applications Gender, age and emotion recognition

Can be applied in following use cases:

- Advertisements The advertisements can be specialized based on the age and the gender of the person on the phone.
- E-commerce Product listings can be altered based on the age and the gender of the person on the phone.
- It also can help identify suspects in criminal cases or at least it can minimize the number of suspects.
- It can provide valuable feedback for in-store promotions by tracking customer emotional responses, enabling retailers to improve product categorization and improve service quality.

LITERATURE SURVEY

The proposed design falls in the CNN framework. A Convolution Neural Network is a deep neural network (DNN) widely used for the purposes of image recognition and processing and NLP. A significant amount of study has been done on human age gender and emotion detection. The major goal of this study is to look at how various Deep learning technologies have been used to determine age, gender, and emotion during the last several years. With the advancement of technology and the application of new methodologies, the accuracy of the outcomes has risen exponentially.

In 2017, Mr. M. Mohammadpour and his team discovered "Facial emotion recognition using deep convolutional networks"[1]. The major goal of this study was to create a model based on facial action units (AUs), which were first recognized by CNN and then used to detect the seven fundamental emotion states. They employed the Cohn-Kanade database to reach the highest accuracy rate of 97.01 by integrating AU, whereas other studies in the literature used a direct CNN and they could only achieve a 95.75% accuracy rate.

Transfer learning using both VGG19 and VGGFace pretrained models by testing the effects of changes in various design schemes and training parameters improved prediction accuracy, according to the paper "Transfer Learning with Deep CNNs for Gender Recognition and Age Estimation" proposed by P. Smith and C. Chen [2]. This research also shows that by using correct model training procedures, high accuracy may be reached.

Mr. Byoung C. published "A Brief Review of Facial Emotion Recognition Based on Visual Information,"[3] which divides approaches into two categories: traditional FER (Facial Emotion Recognition) approaches, which consist of three steps: face and facial component detection, feature extraction, and expression classification; and novel FER (Facial Emotion Recognition) approaches, which consist of four steps: face and facial component detection, feature extraction, and It is believed that FER will be able to enhance its present recognition rate, which includes even spontaneous microexpressions, to that of humans.

Further study discovered that in a work produced by S. Arora and M. P. S. Bhatia, "A Robust Approach for Gender Recognition Using Deep Learning," [4], classification

accuracy of 98.5 percent was reached in just 50 epochs using a mix of convolutional and max pooling layers.

A unique kind of proposition was noticed which was implemented by A. Atanassov and D. Pilev, "Pre-trained Deep Learning Models for Facial Emotions Recognition,"[5]. Pre-trained deep learning CNN models for facial emotion identification were utilized to improve the individualized learning based on their recorded facial expressions. Two pre-trained CNN models for FER were chosen based on these data. One was built on Deep Face CNN, and the other on VGG, both of which improved accuracy by 10% over previous models.

A Research "Deep Learning Classification of Neuro-Emotional Phase Domain Complexity Levels Induced by Affective Video Film Clips," by S. Aydin [6], where PCA is used to PSTM of brief EEG segments, shows that a novel emotional identification approach that proved to detect nine emotional states.

PROBLEM DEFINITION

The main motive is to develop an 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. This implementation proposes a new framework of Convolutional. Neural Network for the simultaneous tasks of age estimation, gender recognition and emotion recognition on face images.

SYSTEM REQUIREMENTS SPECIFICATION

Hardware requirements

- CPU Intel Pentium 2 or higher.
- Cores Minimum dual core.
- Memory 8 GB or higher.
- Hard Disk space 100 GB or higher.

Software requirements

- OpenCV-python
- Tensor flow
- Flask
- Pandas and Numpy
- Keras
- NGROK

SYSTEM DESIGN

The overall high-level architecture of this project is depicted in Figure 4.1. The main motive is to develop an 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. This implementation proposes a new framework of Convolutional Neural Network for the simultaneous tasks of age estimation, gender recognition and emotion recognition on face images.

In this architecture following modules were used:

- TensorFlow: TensorFlow is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. Tensorflow is a symbolic math library based on dataflow and differentiable programming. It is used for both research and production at Google. TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache License 2.0 in 2015.
- Flask: Flask is a micro web framework written in Python. It is classified as a
 microframework because it does not require particular tools or libraries It has
 no database abstraction layer, form validation, or any other components where
 pre-existing third-party libraries provide common functions. However, Flask
 supports extensions that can add application features as if they were
 implemented in Flask itself. Extensions exist for object-relational mappers,
 form validation, upload handling, various open authentication technologies and
 several common framework related tools

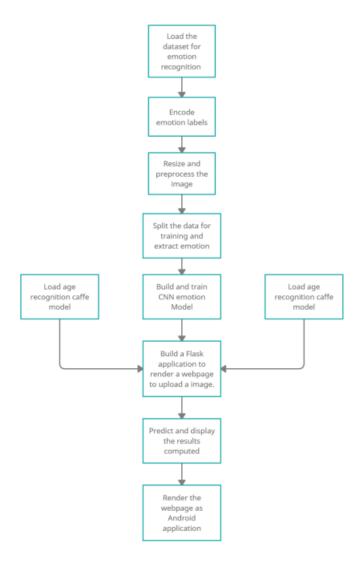


Figure 5.1: System Architecture

- OpenCV: OpenCV (Open-Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel. The library is cross-platform and free for use under the open source Apache 2 License. Starting with 2011, OpenCV features GPU acceleration for real-time operations.
- Caffe: CAFFE (Convolutional Architecture for Fast Feature Embedding) is a
 deep learning framework, originally developed at University of California,
 Berkeley. It is open source, under a BSD license. It is written in C++, with a
 Python interface. Caffe supports many different types of deep learning
 architectures geared towards image classification and image segmentation. It
 supports CNN, RCNN, and LSTM and fully connected neural network designs.

Caffe supports GPU and CPU based acceleration computational kernel libraries such as NVIDIA cuDNN and Intel MKL.

ngrok: ngrok is a tunnelling, reverse proxy that establishes secure tunnels
from a public endpoint to a locally running network service while capturing all
traffic for inspection and replay. Ngrok exposes local servers behind NATs and
firewalls to the public internet over secure tunnels

IMPLEMENTATION

In this section implantation of various modules involved in the design of the Gender, Age and Emotion Recognition using Deep Learning is explored.

Working of Age and gender prediction

Importing cv2, which is an OpenCV library which has the CNN capabilities. Video is captured using the built-in camera. Input can be altered - like usb camera plugged in. Haar cascade model is used to detect faces and eyes in an image. Later, the height and width of the frame is set. Function for loading the caffe models is written. It returns the variable of both the models. These files were downloaded from the Kaggle website. This method was introduced by two Israel researchers, Gil Levi and Tal Hassner in 2015. This contains the information of the trained neural network (trained model). Both. Caffe model files are the model files. The. Pretext files define the layers in the neural network, each layer's inputs, outputs and functionality. The video captured is processed in real time to detect thefaces using the Haar cascade. Once the faces are detected and extracted, this cropped facial image is passed to the age and gender models to make a prediction.

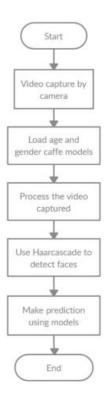


Figure 6.1: Working of Age and Gender Perdition

Working of Emotion prediction

Emotion model is trained using the FER2013. The FER2013 dataset (facial expression recognition) consists of 48*48-pixel grayscale face images. The images are centered and occupy an equal amount of space. This dataset consists of facial emotions of following 7 categories: (angry, disgust, feat, happy, sad, surprise, natural). The dataset is downloaded. Extract it in the data folder with separate train and test directories. Initialize the training and validation generators. Build the convolution network architecture. Finally, compile and train the model. Save the model weights to a ".h5" file. OpenCV Haar cascade xml is used to detect the bounding boxes of face in the webcam and emotion is predicted using the trained model.

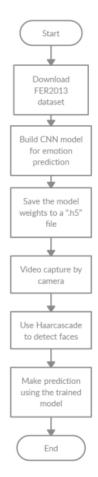


Figure 6.2: Working Of Emotion Prediction

Working of JavaScript component written

Upload button can be used to either select or click an image. Once the image is selected, the "Predict and suggest" button is displayed, clicking it will invoke the computation at the remote workspace and the predicted result is displayed accordingly

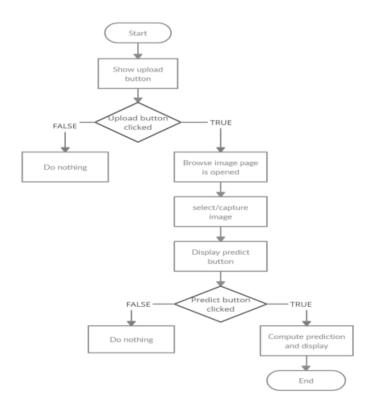


Figure 6.3 Working Of JavaScript Component

Implementation of Flask web application

Flask application is created to render HTML pages and to enable the python code (to make predictions) to communicate seamlessly with the HTML pages. To deploy the web pages for testing purpose NGROK is used. It is a tunneling, reverse proxy that establishes secure tunnels from a public endpoint to a locally running network service while capturing all traffic for inspection and replay. Once the HTML page is rendered, upload button can be used to either select or click an image. Once the image is selected, the "Predict and suggest" button is displayed, clicking it will invoke the computation at the remote workspace and the predicted result is displayed accordingly.

RESULTS

```
n Administrator: C:\Users\User\OneDrive\Desktop\ngrok.exe - ngrok http 80
ngrok
                             adiholla123@gmail.com (Plan: Free)
Account
Version
                             3.2.2
                             India (in)
Region
Latency
                             24ms
Web Interface
                             http://127.0.0.1:4040
                             https://7f68-103-171-59-32.ngrok-free.app -> http://localhost:80
Forwarding
                                                            p50
Connections
                             ttl opn
                                            rt1
                                                   rt5
                                                                    p90
                                                    0.00 0.00
                                             0.00
                                                                    0.00
```

Figure 8.1: NGROK execution Window

Age, Gender and Emotion

Please select the image and upload

Upload/Capture Image...

Figure 8.2 Flask Output Home page

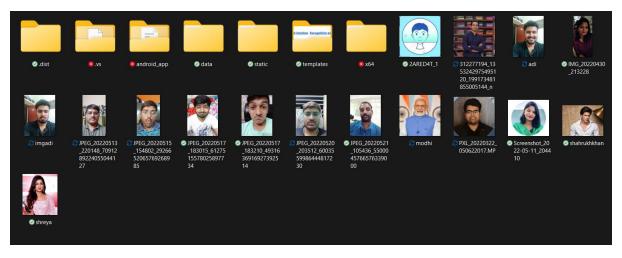


Figure 8.3 Image Browsing Window

Age, Gender and Emotion

Please select the image and upload

Upload/Capture Image...



Predict and Suggest

Figure 8.4 Browsed Image View

Age, Gender and Emotion

Please select the image and upload

Upload/Capture Image...



Age is (21, 28) Gender is Male Emotion is Happy

Figure 8.5 Displaying Predicted Result

```
* ngrok tunnel "NgrokTunnel: "http://29ea-103-171-59-35.ngrok.io" -> "http://localhost:
5000"" -> "http://127.0.0.1:5000"
 * Serving Flask app "app" (lazy loading)
 * Environment: production
   Use a production WSGI server instead.
 * Debug mode: off
 * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
127.0.0.1 - - [10/Apr/2023 19:29:18] "GET / HTTP/1.1" 200
127.0.0.1 - - [10/Apr/2023 19:29:19] "GET /static/css/main.css HTTP/1.1" 304 -
127.0.0.1 - - [10/Apr/2023 19:29:19] "GET /static/js/main.js HTTP/1.1" 304 -
127.0.0.1 - - [10/Apr/2023 19:29:20] "GET
                                          icon.ico HTTP/1.1" 404 -
1/1 [=======] - 2s 2s/step
Fearful
127.0.0.1 - - [10/Apr/2023 19:32:12] "POST /predict HTTP/1.1" 200 -
Male (21, 28)
1/1 [======] - 0s 61ms/step
127.0.0.1 - - [10/Apr/2023 19:32:21] "POST /predict HTTP/1.1" 200 -
Нарру
Male (15, 20)
1/1 [======= ] - 0s 63ms/step
127.0.0.1 - - [10/Apr/2023 19:32:35] "POST /predict HTTP/1.1" 200 -
Surprised
Male (21, 28)
1/1 [======
              127.0.0.1 - - [10/Apr/2023 19:32:43] "POST /predict HTTP/1.1" 200 -
Fearful
Male (21.
```

Figure 8.6: Python Console after Prediction

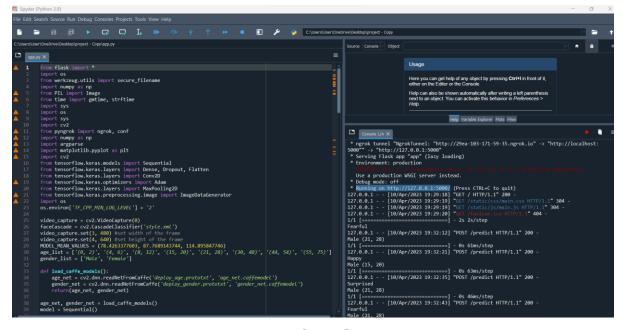


Figure 8.6 Code Snippets

CONCLUSION AND FUTURE WORK

The proposed framework is capable of emotion recognition, age estimation and gender recognition simultaneously in a fast and efficient way. As a part of outcome, and a flask application is developed to deploy web application. Misclassifications might occur because people from different ethnic groups have different facial characters, which might slightly alter the actual age from the predicted age. The use of glasses might affect the emotion classification by interfering with the features learned. The models used in our research have been using the sigmoid function, our implementation is much better because we have used a ReLU function to do the same which increases its speed and improves accuracy. When we compare our project and implementation with the other implementations mentioned in our research, we can see a significant improvement in accuracy for detecting the age and gender. Additionally, our application combines age and gender with emotion to create personalized results catering to the respective age and emotion. Our model has an effective accuracy of 93% for age, 80% for gender and 85% for emotion. Finally, it would be interesting to implement this model on a live business service. This project can be applied in social media to customize advertisements and content, in ecommerce to customize product listing, in t h e entertainment industry to customize OTT (over the top) content and finally to profile criminals and help catch them either in the act or when they have escaped

REFERENCES

- [1] M. Mohammadpour, H. Khaliliardali, S. M. R. Hashemi and M. M. AlyanNezhadi, "Facial emotion recognition using deep convolutional networks," 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), 2017, pp. 0017-0021, doi: 10.1109/KBEI.2017.8324974.
- [2] P. Smith and C. Chen, "Transfer Learning with Deep CNNs for Gender Recognition and Age Estimation," 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 2564-2571, doi: 10.1109/BigData.2018.8621891.
- [3] Ko, Byoung C. 2018. "A Brief Review of Facial Emotion Recognition Based on Visual Information" Sensors 18, no. 2: 401.
- [4] S. Arora and M. P. S. Bhatia, "A Robust Approach for Gender Recognition Using Deep Learning," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1-6, doi: 10.1109/ICCCNT.2018.8494194.
- [5] A. Atanassov and D. Pilev, "Pre-trained Deep Learning Models for Facial Emotions Recognition," 2020 International Conference Automatics and Informatics (ICAI), 2020, pp. 1-6, doi: 10.1109/ICAI50593.2020.9311334.
- [6.] S. Aydın, "Deep Learning Classification of Neuro-Emotional Phase Domain Complexity Levels Induced by Affective Video Film Clips," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 6, pp. 1695-1702, June 2020, doi: 10.1109/JBHI.2019.2959843.
- [7] M. M. Taghi Zadeh, M. Imani and B. Majidi, "Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters," 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI), 2019, pp. 577-581, doi: 10.1109/KBEI.2019.8734943.
- [8] P. Aiswarya, Manish and P. Mangalraj, "Emotion recognition by inclusion of age and gender parameters with a novel hierarchical approach using deep learning," 2020 Advanced Communication Technologies and Signal Processing (ACTS), 2020, pp. 1-6, doi: 10.1109/ACTS49415.2020.9350479.
- [9] I. Lasri, A. R. Solh and M. E. Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network," 2019 Third International Conference on

Intelligent Computing in Data Sciences (ICDS), 2019, pp. 1-6, doi: 10.1109/ICDS47004.2019.8942386.

[10] R. Helaly, M. A. Hajjaji, F. M'Sahli and A. Mtibaa, "Deep Convolution Neural Network Implementation for Emotion Recognition System," 2020 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2020, pp. 261-265, doi: 10.1109/STA50679.2020.9329302.