Facial Emotion Detection using Deep Learning for Psychometric Assessment in a Cloud Environment

MSc Research Project
MSc. in Cloud Computing

Sanal Sudhakaran
Student ID: 21125775

School of Computing
National College of Ireland

Supervisor: Dr. Rashid Mijumbi

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Sanal Sudhakaran
Student ID:	21125775
Programme:	MSc. in Cloud Computing
Year:	2022
Module:	MSc Research Project
Supervisor:	Dr. Rashid Mijumbi
Submission Due Date:	15/12/2022
Project Title:	Facial Emotion Detection using Deep Learning for Psychometric Assessment in a Cloud Environment
Word Count:	5665
Page Count:	18

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Sanal Sudhakaran
Date:	31st January 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Facial Emotion Detection using Deep Learning for Psychometric Assessment in a Cloud Environment

Sanal Sudhakaran
21125775

Abstract

One of the most effective ways to convey feeling is through one's face. People's facial expressions can convey a wide range of feelings, from joy to sorrow to anger to surprise to disgust. The facial emotion recognition system are adopted by many industries such as narcotics, esports, hospitals and more. Individuals' mental, emotional, and behavioral faculties can be evaluated via psychometric assessments, which are widely adopted by businesses because of Covid-19 pandemic. Personality, skills, and passions can all be assessed with their help. Problem-solving, memorization, verbal comprehension, focus, and reaction time are some of the abilities typically measured by these tests. An individual's strengths, weaknesses, and development opportunities can all be revealed with their help. Personality, skills, and passions can all be assessed with their help. Problem-solving, memorization, verbal comprehension, focus, and reaction time are some of the abilities typically measured by these tests. An individual's strengths, weaknesses, and development opportunities can all be revealed with their help. In this study a facial emotion recognition system is developed using CNN architecture for psychometric assessments. The models that are used for evaluation are two Custom CNN based models consisting of 3 and 5 layers which will be implemented for Psychometric assessments.

1 Introduction

1.1 Background

Emotion analysis of human faces is the focus of facial emotion detection, a technology that employs computer vision algorithms. One can utilize it to label a range of feelings, from joy to shock to anger to fear to sadness. Research, advertising, instruction, and treatment are just some of the many applications for the technology. It finds additional applications in the realms of gaming, VR, and robotics. To better comprehend one's own and other people's emotional states, facial emotion detection can be put to good use. Furthermore, it can be used to report on people's reactions to events.

Assessments of an individual's intelligence, character, and other psychological traits can be obtained through the use of psychometric tests. These kinds of tests see widespread use in a range of fields, from academia and the workplace to psychiatry and substance abuse treatment. There has been a rise in the use of psychometric tests in recent years, especially during the COVID-19 pandemic, as more businesses adopt remote staffing and online application and screening procedures. These tests are helpful for assessing candidates and employee's mental capacities and character traits in an online setting because they can be given without the need for face-to-face contactOlson-Buchanan et al. (2013). In addition, psychometric assessments have become increasingly utilized in the field of mental health care for the purpose of diagnosis and treatment. As more and more businesses adopt remote work and digital hiring procedures, the prevalence of psychometric tests is expected to riseBorsboom et al. (2013).

Facial emotion detection makes extensive use of machine learning and deep learning. The shape of a person's eyes, mouth, and eyebrows, among other features, are analyzed by machine learning algorithms to deduce their emotional state. In order to recognize and classify various emotions, these algorithms are typically trained on large datasets of images of people's faces, along with labels indicating what emotion they are feeling. Contrarily, deep learning algorithms are a form of machine learning algorithms intended to simulate human cognitive processes. These algorithms are typically composed of many nested layers of interconnected nodes that perform hierarchical processing and analysis of the input data in order to learn and recognize intricate patterns and relationshipsMahata and Phadikar (2017). A common practice in the field of facial emotion detection is to use deep learning algorithms to examine pictures of people's faces and deduce how they feel. Facial emotion detection systems that employ machine learning and deep learning techniques are highly accurate and robust, making them useful in many contextsConnie et al. (2017).

Together, psychometric testing and facial emotion detection technology can shed light on a person's unique blend of intelligible traits. A person's mental state and assessment responses can be gleaned from their facial expressions using facial emotion detection technology, for instance. This can be done by analyzing the assessment taker's expressions for signs of anxiety, frustration, or distraction. Mental health professionals and other evaluators can use this data to form a more nuanced picture of the patient's personality and mental state and thus make more accurate diagnosis and treatment recommendations. Additionally, the results of a psychometric assessment may be corroborated with the help of facial emotion detection technology by providing additional proof of the subject's emotional state at the time of the assessmentOlson-Buchanan et al. (2013). As a whole, incorporating facial emotion detection technology into psychometric assessments has the potential to yield more precise insights into an individual's mental capacities and personality traits.

In this research project, emotion detection system is developed which will anticipate the emotion of the human using the facial expressions. A deep learning approach is used in order to develop the models. The dataset used is Emotion detection dataset which is obtained from Kaggle. Various various deep learning techniques are used convert the raw data into a dataset which is used to train and validate the models. The models used for evaluation purpose are CNN based models, two custom CNN models with 3 and 5 layers respectively are developed and compared to select the most suitable one. The models classifies the human emotion into seven different classes, namely, happy, sad, neutral, fearful, angry, disgusted and surprised will be used. The results and evaluation obtained from the experiments will assists in psychometric assessments which evaluates subjects based on their emotional and cognitive reactions to different scenarios and questionnaires.

1.2 Reserach Question

How can deep learning techniques(CNN) help in detecting human emotion using facial expressions and categorizing into seven different classes (happy, sad, neutral, angry, surprised, disgust and fearful), which will assist in enhancing psychometric reports during online assessments?

2 Related Work

Facial emotion detection has been implemented using standard machine learning methods like Support Vector Machines, Random Forests, and Decision Trees. These strategies for emotion recognition rely on features of the face, such as its shape and texture, that must be extracted manually.Deep learning methods are widely adopted to recognize human emotions in facial expressions. These techniques take facial photographs as input and automatically extract features from them to study a variety of emotions. Furthermore, recurrent neural networks (RNNs) can learn the temporal changes in facial expressions and apply this knowledge to the task of emotion detection in video.This section of the paper reviews the state-of-the-art studies on emotion recognition in faces, covering both traditional machine learning and deep learning approaches. It also consists of a review on Psychometric Assessments

2.1 Conventional Machine Learning Techniques for Facial Emotion Detection

Varma et al. (2020) This technical report examined the use of PCA and LDA in facial feature extraction for an emotional recognition system. The results of the experiment concluded that LDA outperformed PCA on its own, and that combining the features of PCA and LDA further improved the system's performance. Additionally, the SVM classifier was found to outperform the HMM classifier across all three feature sets. This research demonstrates that PCA and LDA can be effectively used together to accurately detect and interpret people's emotions. Reddy et al. (2019)In this study, a new approach for recognizing facial expressions in images was presented. It involves combining local and global features extracted from the input images using Gabor and Haar wavelets, respectively, and then reducing their dimensionality using nonlinear principal component analysis. The combined features were fused and fed into a support vector machine classifier to identify six emotions (happiness, surprise, fear, disgust, anger, and sadness). The Extended Cohn-Kanade (CK+) database was used to evaluate the performance of the proposed method, which was found to be more effective than using a single set of features.

Expressions on people's faces can convey a lot about their personalities and habits. Scientists in the field of psychology have discovered that the eyes, nose, and mouth's relative positions are key indicators in the process of identifying an individual's emotional state. Distance metrics, which employed high-dimensional rate transformation and regional volumetric difference maps to classify and quantify facial expressions, were among the earliest approaches to estimating the intensity of facial emotions. Principal component analysis (PCA) is widely used in video analysis systems to represent facial expressions. Facial action units responsible for conveying various emotions have also been isolated with PCA Cornejo et al. (2015). To create a facial action unit, PCA is used in other methods of facial expression recognition Mahata and Phadikar (2017). Using the active contour model, Siddiqi et al. (2015) successfully isolated the face from its surrounding environment, and then applied the energy functions developed to minimize differences within the face while maximizing the distance to the surrounding environment. The researchers also used wavelet decomposition to cut down on noise, and they used optical flow to extract geometric features of facial expressions and expression movements. Embedded devices often lack the processing power and memory that are necessary for implementing traditional machine learning techniques . Therefore, it is essential to think about techniques that can perform classification in real time using little computational power while still yielding satisfactory results.

2.2 Deep Learning Techniques for Facial Recognition System

In order to identify feelings conveyed in videos, Li et al. (2019) proposed a three-dimensional convolutional neural network (CNN) architecture. They used the CASME II, CASME, and SMIC benchmark datasets to evaluate the deep features they extracted. Face cropping and rotation techniques for feature extraction using a CNN were also performed by Li et al. (2013), and their method was tested on the CK+ and JAFFE databases. To deal with occlusions in the face and concrete on more discriminative, non-occluded regions, Li et al. (2018) proposed the adaptive convolutional neural network (ACNN), another CNN-based approach. The ACNN employs a full-stack learning system, which first combines multiple facial ROI representations into one, and then assigns weights to those representations using a proposed gate unit that determines an adaptive weight based on the importance of the area. Patch-based CNN (pACNN) and global-local-based ACNN are two variations of the network designed to work with varying regions of interest (gACNN). Three distinct CNN architectures were used by Lopes et al. (2017). to categorize human facial expressions into distinct emotional categories. Breuer and Kimmel (2017) used multiple FER datasets to train a model that could identify seven core emotions. A multi-angle optimal pattern-dependent deep learning (MAOP-DL) system was proposed by Zhang et al. (2017) and Jain et al. (2020) to deal with sudden shifts in lighting and to align the feature set via optimal arrangements based on multiple angles. Their method starts by erasing the background, then separates the subject from the face images, and finally eliminates the texture patterns and primary features of the facial landmarks. For facial expression prediction, the extracted features are fed into a long short-term memory convolutional neural network.

When building a classification model for recognizing facial expressions, Connie et al. (2017) combined the use of convolutional neural networks (CNNs) with the use of scale-invariant feature transform (SIFT) features. The system proposed by Jung et al. (2015) makes use of two distinct Convolutional Neural Network (CNN) models. The first model used facial landmarks to determine the subject's presence, while the second used temporal geometry to determine the subject's age. This novel integration scheme combines these two models in an effort to boost FER's efficiency. In 2015, Yu and Zhang (2015) used a hybrid CNN to achieve state-of-the-art results in Facial Emotion Recognition. During training, they applied transformations to the input image and used a network of convolutional neural networks (CNNs) with five convolution layers for each facial expression. Their model was able to predict several feelings for each test subject. For maximum efficiency, they opted for stochastic pooling rather than max pooling. By fusing an LSTM with a DL-based visual feature extractor like a CNN model, these hybrid techniques can recognize emotions from image sequences.

2.3 Review on Psychometric Assessment

Psychological or psychometric assessment is a method of gauging one's mental faculties through tests of personality, intelligence, and other traits. It's a vital resource for figuring out what drives and motivates a person and how they act. Intelligence, personality, and aptitude are just some of the psychological traits that can be measured with this toolBorsboom et al. (2013).

For a long time, people have relied on psychometric testing to help them learn more about themselves and the people around them. It has been applied to the assessment of intelligence, aptitude, and personality, among other areas of mental health. It's also a standard way to evaluate potential employees and foretell how they'll do on the job.Much has been written about psychometric assessment, with numerous studies documenting the tests' validity and reliability in addition to their efficacy in a variety of contexts. Psychometric tests have been shown to be valid and reliable, making them an important resource for studying human behavior and character. However, the use of psychometric assessment is fraught with ethical concerns, especially in the workplace.When it comes to learning about people and their actions, psychometric testing is invaluable. According to the studies conducted on it, it is a valid and reliable method of evaluating one's mental health, and can be used in the detection of mental disorders and the forecasting of one's productivity at workOlson-Buchanan et al. (2013).

2.4 Summary of the Related Work

Authors	Method	No. of Emotions Analyzed
Varma et al. (2020)	SVM and HMM	6
Li et al. (2018)	CNN and ANN	7
Breuer and Kimmel (2017)	CNN with transfer learning	7
Zhang et al. (2017)	PHRNN and MSCNN	7
Jung et al. (2015)	DTAN and DTGN	7

Table 1: Summary of the Literature Review

Traditional machine learning methods such as SVM, HMM, and the KNN all performed admirably. The field of emotion detection using facial expressions has seen widespread adoption of Deep Learning techniques like Convolutional Neural Networks and Artificial Neural Networks. In this study, we will use convolutional neural network (CNN) models, two custom CNN models with 3 and 5 layers to build a facial emotion recognition system for use in psychometric testing, where it will be used to assess the candidate's emotional and cognitive reaction to different scenarios and questions.

3 Methodology

Facial emotion detection technology analyzes a person's facial expressions to determine what emotion they are experiencing. This technology has numerous applications, including mental health care, law enforcement, and customer service. Researchers from all over the world use a variety of traditional machine learning and deep learning techniques to develop facial emotion recognition systems. In this study, a deep learning-based framework is used to develop a facial emotion recognition system that will aid in psychometric-based assessment by interpreting the subject's emotional response to each question. The research is planned using a modified Cross-Industry Standard Process for Data Mining (CRISP-DM). As can be seen in Figure 1, the modified CRISP-DM model was developed for use in facial emotion detection for psychometric assessments.

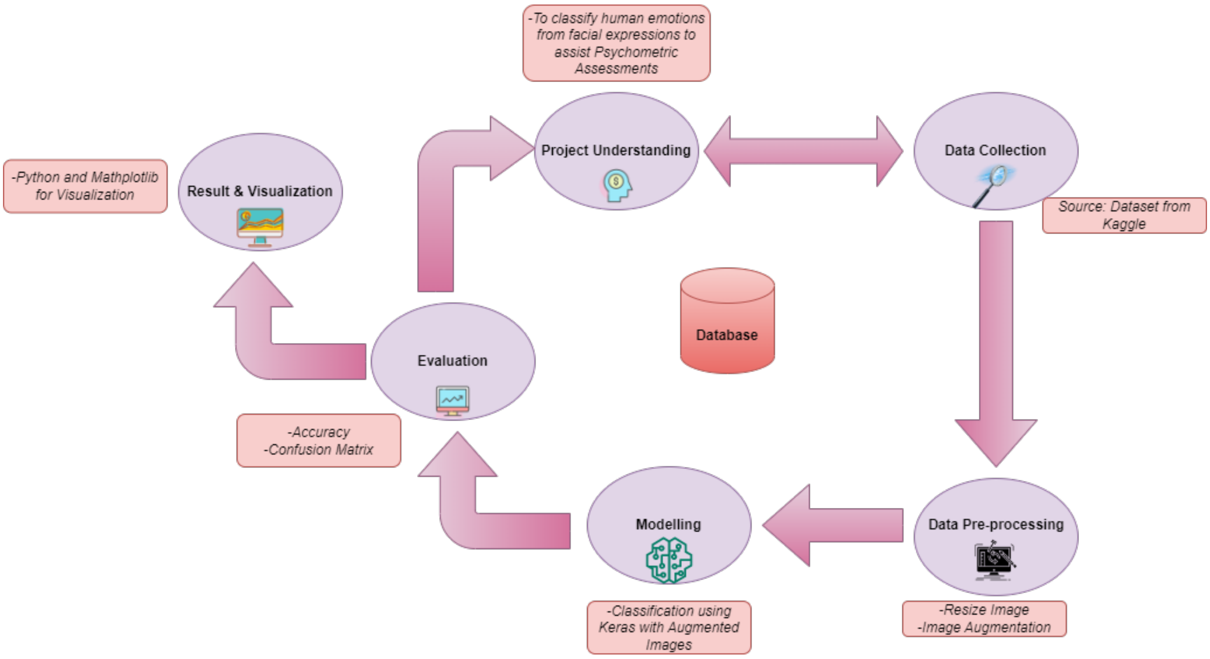


Figure 1: Emotion Detection Methodology using Modified CRISP-DM

- **Project Understanding:**To better evaluate psychometric tests, the project's understanding phase outlines how facial expressions can be used to detect human emotion and be used psychometric assessments.
- **Data Collection:** Kaggle, a publicly available platform, was used for the dataset. There were 35,685 face images in grayscale (48x48 pixels) used from the Emotion Detection dataset, which was split into a training set and a testing set. Pictures are sorted into happy, neutral, sad, angry, surprised, disgusted, and fearful categories based on the feelings conveyed by the subjects' facial expressions.
- **Data Pre-Processing:**To prepare the acquired images for further analysis, we employ a number of data augmentation techniques on the raw image data. Among the data augmentation techniques are operations like flipping a picture horizontally, resizing it, varying the contrast, etc. Improving model performance is facilitated by the data pre-processing task. The features have been extracted from the images in an effective manner by using the pre-trained weights of ImageNet.
- **Modelling:** Here we use various deep learning models for classification strategies. In total, the research project uses two models to devise a facial emotion recognition system: Two custom CNN models of varying layer counts such as 3 and 5.
- **Evaluation::** Here, we gather a variety of metrics to use in assessing the performance of the deep learning models employed.

- **Results & Visualization:**In this last phase of research project, it is determined whether the project objectives meets the desire outcomes or not, along with project goals and results.

4 Design Specification

The human facial emotion detection project's architecture has two tiers "Client Tier" and the "Business Logic Tier," as shown in Figure 2 . In the client layer, Matplotlib is used to graphically represent the results of the classification models and exploratory data analysis is done using Keras(python library). Data collection, data pre-processing, data augmentation, modeling, deployment, training, and evaluation of classification models all take place in the business logic tier.

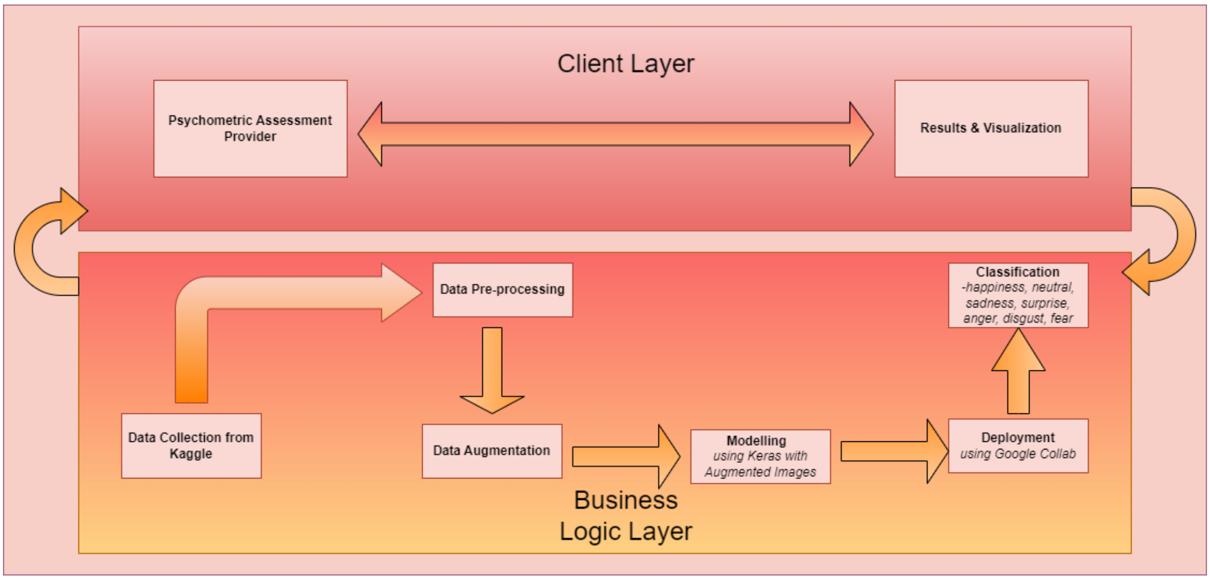


Figure 2: Two Tier Architecture for the System

Since deep learning and convolutional neural networks (CNN) can automatically learn features from raw data and achieve high accuracy in facial emotion recognition, they are frequently used for this task. CNNs can also be used to recognize emotions that are not readily distinguishable from one another and to identify patterns in facial expressions that are indicative of certain emotions. CNNs are especially useful for video surveillance and automated facial recognition because of their ability to recognize facial expressions from a variety of anglesJain et al. (2020). In this project CNN based models are used for analyzing the dataset and predicating the human emotion detection using the facial expression for psychometric assessment. The models used are two custom CNN models consisting of 3 and 5 layers.

- **Custom CNN Model consisting of 3 Layers:** The model is a three-layered custom sequential model . The first layer consist of 64-neuron, 2-dimensional convolutional layer using the same padding and a (3, 3) filter. Next comes a maximum pool layer where the pool size is fixed at (2, 2). The second layer is just like the first, except it has 128 neurons and a (3, 3) filter while maintaining the same padding.

The next layer is a maximum pool layer, and its pool size is (2, 2). The third layer is the same as the first two, except that this time there are 256 neurons and the filter is (3, 3), and the padding is also the same. The next layer is a maximum pool layer, with a pool size of (2, 2). Each of the model's three layers has "relu" entered as its activation keyword.

- **Custom CNN Model consisting of 5 Layers:** The first layer of this 5-layer sequential model consists of 32 neurons in a 2D convolutional layer using a (3, 3) filter with the same-valued padding. In order to activate it, the "relu" activation keyword has been entered. Next comes a maximum pool layer where the pool size is fixed at (2, 2). The second layer is made up of 64 neurons and uses a 3x3 convolutional filter with the same padding as the first layer. The "relu" activation keyword has been applied. Next comes a maximum pool layer where the pool size is fixed at (2, 2). In this case, we have also included a dropout function with a 0.5-valued parameter and a batch normalization function. Third layer is identical to second except it has 128 neurons instead of 64. The structure of the first layer is repeated in the fourth and fifth layers, with 256 and 512 neurons . Then a flatten() function is utilized, followed by a dropout function with a 0.5 value passed for its parameter, to remove the final two layers of the network. Moreover, the last layer model is activated by a soft max function.

5 Implementation

In Figure3 the workflow of the facial emotion detection system which is used for psychometric assessments is illustrated.

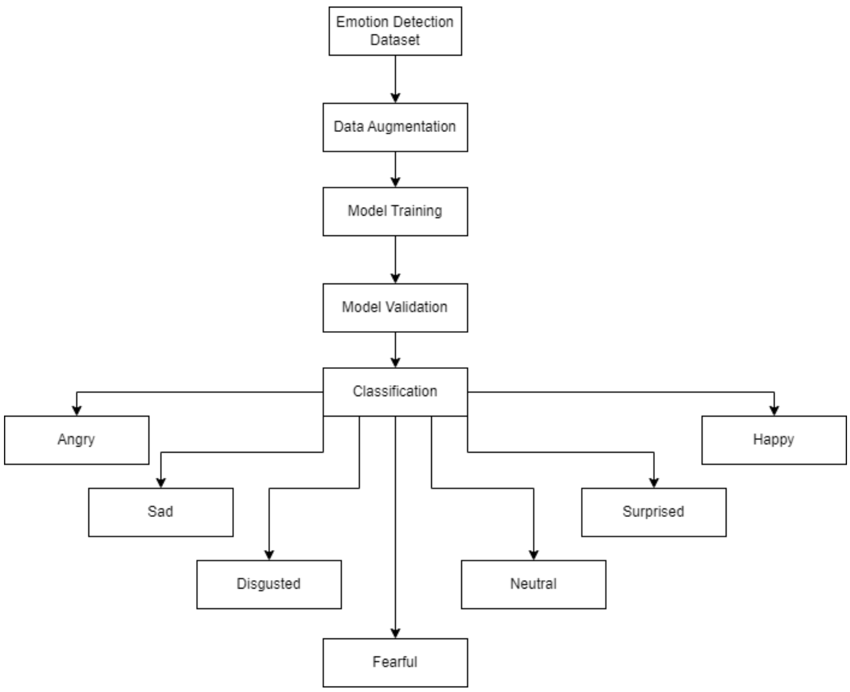


Figure 3: Workflow of the Facial Emotion Detection Model

The dataset used is Emotion detection dataset which consists of 35,685 samples, which is used as a 48x48 pixel gray scale images of faces which is then split into 80% training data and 20% test data. The test data consists of 958 samples for angry, 111 disgusted, 1024 fearful, 1774 happy, 1233 neutral, 831 surprised and 1247 sad images. On the other the training data consists of 3995 angry images for disgusted, 436 disgusted, 4097 fearful, 7215 happy, 4986 neutral , 4830 sad and 3171 samples for surprised. ADAM is used as the optimization function, and the categorical cross entropy is used as the loss function, in the implementation of the models. Several libraries, such as pandas, numpy, matplotlib, seaborn, tensorflow, keras, and opencv, were used to for the research work.

In data pre-processing, in order to avoid underfittng and over-fitting, data augmentation techniques are used. In machine learning, the problem of making too specific or too general of predictions is referred to as overfitting or underfitting, respectively. When a model is overly complicated and considers too many variables, it overfits the data used to train it and produces predictions that are useless when applied to new data. The problem of underfitting arises when a model is oversimplified and fails to account for enough of the relevant variables, leading to predictions that are too broad to be grounded in the actual data Li et al. (2013). The data augmentation techniques used are image rescaling, width shift, height shift, horizontal flip and zoom range for training and validation as shown in Figure 4

```
#preparing training data
img_size=48
train_datagen = ImageDataGenerator(      width_shift_range = 0.1,
                                         height_shift_range = 0.1,
                                         horizontal_flip = True,
                                         rescale = 1./255,
                                         validation_split = 0.2
                                        )

train_generator = train_datagen.flow_from_directory(directory = train_dir,
                                                    target_size = (img_size,img_size),
                                                    batch_size = 64,
                                                    color_mode = "grayscale",
                                                    class_mode = "categorical",
                                                    subset = "training"
                                                   )
```

```
Found 22968 images belonging to 7 classes.
```

```
#preparing validation dataset
validation_datagen = ImageDataGenerator(rescale = 1./255,
                                         validation_split = 0.2)

validation_generator = validation_datagen.flow_from_directory( directory = test_dir,
                                                              target_size = (img_size,img_size),
                                                              batch_size = 64,
                                                              color_mode = "grayscale",
                                                              class_mode = "categorical",
                                                              subset = "validation"
                                                             )
```

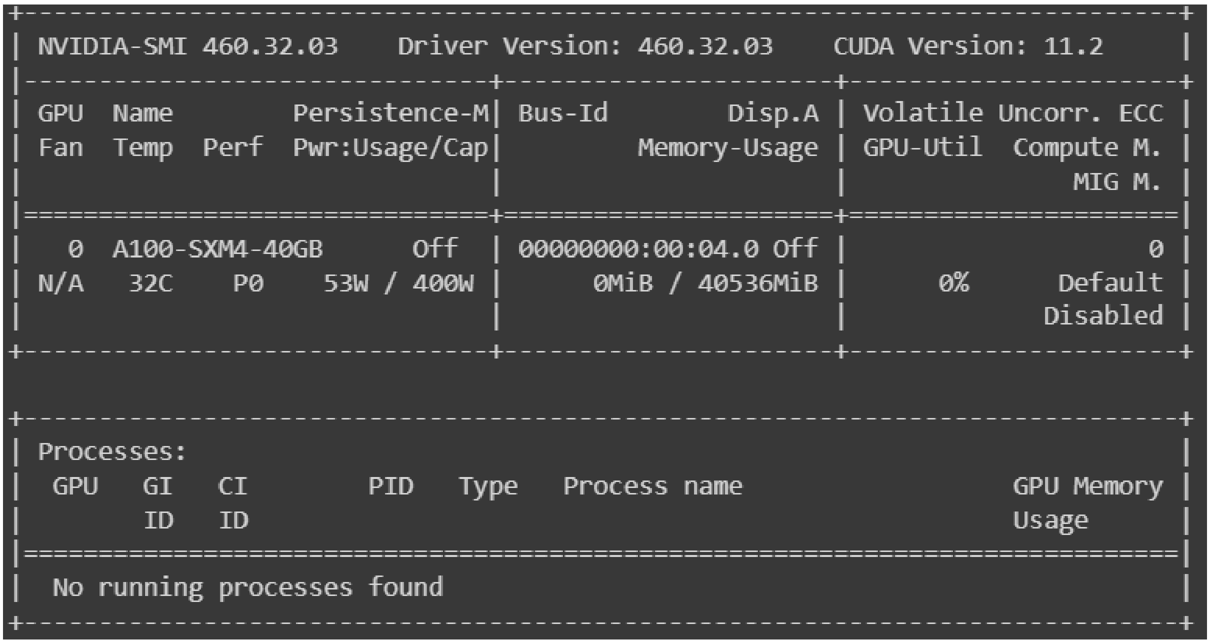
```
Found 1432 images belonging to 7 classes.
```

Figure 4: Data Pre-processing for training and validation

After the models have been trained, they have been visualized using the matplotlib and pandas libraries. The entire experiment is implemented using Python and run on Google's Colab platform. It is possible to detect and classify seven distinct emotions from a person's face: happiness, sadness, neutrality, anger, disgust, fear, and surprise as shown in Figure 5. As per person's mental state, assessment responses can be derived from their facial expressions using facial emotion detection technology, this can aid psychometric assessments.

Figure 5: Classes of 7 Human Emotions

The developed framework can be deployed either in the cloud or on a local machine. Given that the system requires image processing for developing the models, the taring of the model utilizes graphical processing unit (GPU). To train, validate and evaluate and visualize the results Google Colab was used . The best model for the facial emotion detection system can then be used for psychometric assessment, which aids in evaluate test subjects according to their mental and emotional response to the questionnaire. The Specs are shown in the below Figure6



```
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 460.32.03    Driver Version: 460.32.03    CUDA Version: 11.2     |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|                               |                      |               MIG M. |
|===============================+======================+======================|
|   0  A100-SXM4-40GB      Off  | 00000000:00:04.0 Off |                    0 |
| N/A   32C    P0    53W / 400W |      0MiB / 40536MiB |      0%      Default |
|                               |                      |             Disabled |
+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Processes:                                                                  |
|  GPU   GI   CI        PID   Type   Process name                  GPU Memory |
|        ID   ID                                                   Usage      |
|=============================================================================|
|  No running processes found                                                 |
+-----------------------------------------------------------------------------+
```

Figure 6: Specification of the Google Colab Environment

6 Evaluation & Results

Each facial emotion detection algorithm must be ranked according to a set of criteria in order to determine which one is best. In order to enhance the classification accuracy, this study employs a custom deep learning architecture based on a convolutional neural network, with both a 3-layer and 5-layer structure. There will be a comparison and evaluation of these models using certain metrics. These models, has been trained for 45 and 20 epochs respectively.

6.1 Experiment / Evaluation of Custom 3-layered CNN Model

To achieve a better performance metric, a three-layer Custom Convolution neural network is developed, with the activation function used as relu for each convolutional layer. Comparatively, we use signmoid and the softmax function in conjunction with a two-layer densenet. This model, has been trained for 45 epochs. In the results section, outcomes achieved from the CNN-based Custom Model (3 layers) are shown.

6.2 Experiment / Evaluation of Custom 5-layered CNN Model

To further enhance the detection rate, a new custom model has been developed. A five-layer convolutional neural network is used in this model. For the convolutional layers, the relu activation function used, while the dense layers make use of the sigmoid and softmax functions. Within 20 epochs, the model will be trained.

6.3 Results

Training accuracy, validation accuracy, training loss, and validation loss are utilized to determine the overall success of each individual Custom CNN based (3 and 5 layered) model. The results are in a form graphical representation.

- **Accuracy:**The accuracy of a model can be determined by comparing the fraction of correct predictions to the total number of observations. The accuracy reveals how well the model functions as a whole.
- **Recall:**The precision of a model is measured by its precision score, which provides information about the prevalence of false positives in the dataset. Having fewer false positives corresponds to a higher precision score.
- **F1-Score:** When assessing the efficacy of a model, the recall score is most useful because it reveals the presence of false negatives in the dataset. A lower number of false negatives corresponds to a higher Recall score.

$$f1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{1}$$

- **Support:**To what extent a particular class represents the true response is what is meant by "support" . It is the proportion of data points where the actual response matches the model's prediction.
- **Confusion Matrix:**The effectiveness of a classification system can be measured with a table called a confusion matrix. It's a table that summarizes a classification model's efficacy by contrasting observed and predicted data. True positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) are the four possible binary classification outcomes that make up the confusion matrix (FN). The number of times a given classification outcome was predicted is displayed in each result. A classification model's accuracy can be measured, and error sources can be pinpointed, with the help of the confusion matrix.

6.3.1 Results for Custom CNN Model(3 Layers)

The train accuracy for Custom CNN Model (3-layers) is 61.42% on the other hand the validation accuracy 62.01%. The graphical representation for the training loss and validation loss, and training accuracy and validation accuracy has been illustrated in Figure7.

Figure8. demonstrates the confusion matrix of Custom CNN Model(3-layers).

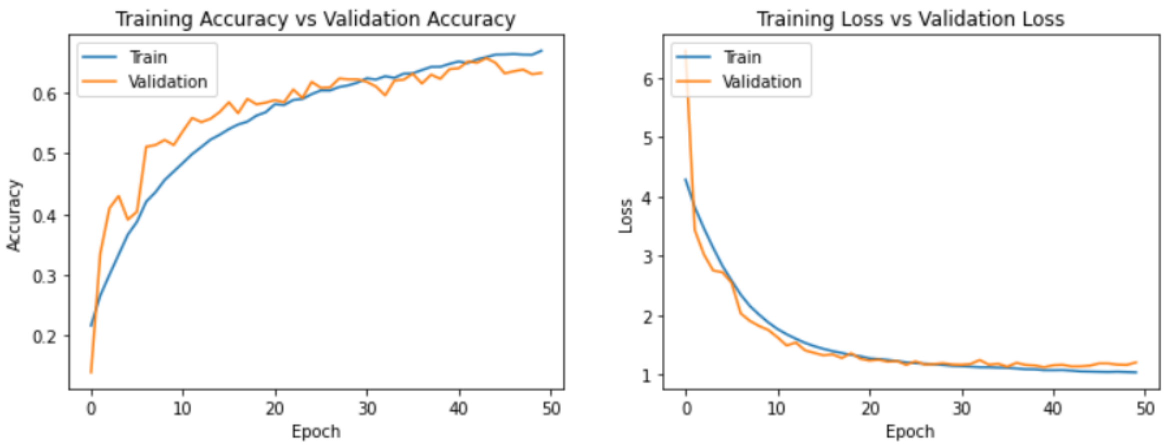


Figure 7: Graphical Representation of Accuracy and Loss for Custom CNN Model(3 Layers)

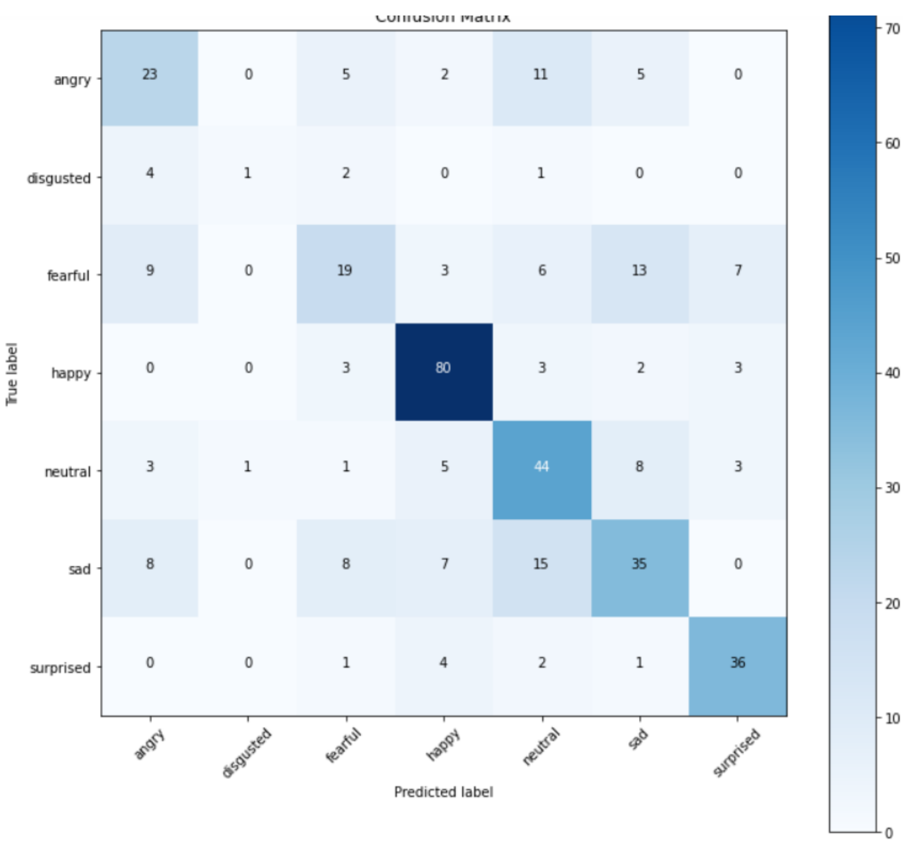


Figure 8: Confusion Matrix of Custom CNN Model(3-Layers)

The classification report is represented in Figure9.

Classification Report

	precision	recall	f1-score	support
angry	0.11	0.12	0.12	191
disgusted	0.00	0.00	0.00	22
fearful	0.13	0.09	0.11	204
happy	0.28	0.31	0.29	354
neutral	0.19	0.22	0.21	246
sad	0.17	0.15	0.16	249
surprised	0.11	0.13	0.12	166
accuracy			0.18	1432
macro avg	0.14	0.15	0.14	1432
weighted avg	0.18	0.18	0.18	1432

Figure 9: Classification Report of Custom CNN Model(3-Layers)

6.3.2 Results for Custom CNN Model(5 Layers)

The train accuracy for Custom CNN Model (5-layers) is 63.75% . The graphical representation for the training loss training accuracy and validation accuracy has been illustrated in Figure10.

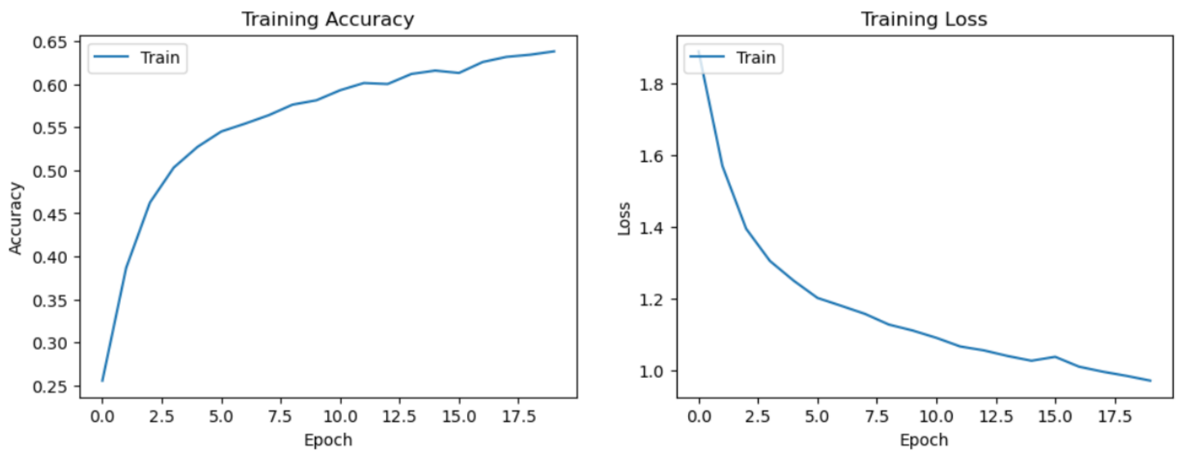


Figure 10: Graphical Representation of Accuracy and Loss for Custom CNN Model(5 Layers)

Figure11. demonstrates the confusion matrix of Custom CNN Model(5-layers).

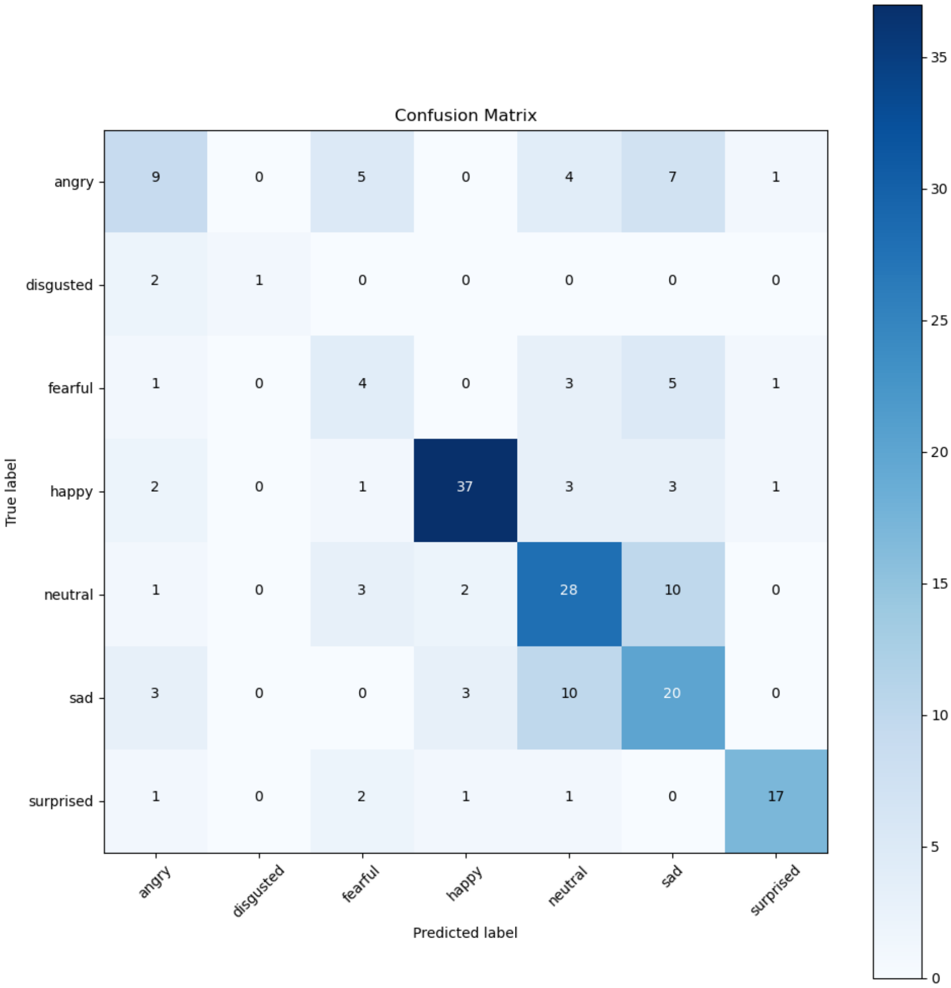


Figure 11: Confusion Matrix of Custom CNN Model(5-Layers)

The classification report is represented in Figure12.

Classification Report

	precision	recall	f1-score	support
angry	0.13	0.14	0.13	958
disgusted	0.02	0.01	0.01	111
fearful	0.14	0.10	0.11	1024
happy	0.25	0.22	0.23	1774
neutral	0.16	0.17	0.16	1233
sad	0.18	0.24	0.21	1247
surprised	0.13	0.12	0.13	831
accuracy			0.17	7178
macro avg	0.14	0.14	0.14	7178
weighted avg	0.17	0.17	0.17	7178

Figure 12: Classification Report of Custom CNN Model(5-Layers)

6.4 Discussion

Emotion detection using facial expressions was implemented using Custom-CNN models of 3 and 5 layers respectively. Both models successfully analyzed and categorized the human emotions into 7 different classes such as sad, angry, happy, neutral, surprised, disgusted and fearful and the the results are evaluated. It is observed that the train accuracy for Custom CNN Model (3-layers) is 61.42% on the other hand the validation accuracy 62.01%.Train accuracy for Custom CNN Model (5-layers) is 63.75% . The classification of the human emotions will aid in assessing the candidate's emotional and cognitive reaction to different scenarios and questions.

7 Conclusion and Future Work

Challenged as it may be, emotion detection based on facial expressions remains a rapidly growing field for study. To aid in psychometric evaluations, this study developed two distinct deep learning custom CNN models capable of accurately identifying and predicting human emotion based on facial expressions, and then further classifying those expressions into seven distinct categories: happy, sad, neutral, surprised, disgusted, fearful, and angry. Facial emotion recognition is a vast and rapidly developing area of study, with researchers using a variety of deep learning models and algorithms to achieve high levels of accuracy.

Facial expression-based emotion detection was implemented using 3- and 5-layer Custom-CNN models. Both models are assessed based on their ability to analyze and classify human emotions into seven distinct categories, including happy, sad, angry, neutral, surprised, disgusted, and fearful. While the validation accuracy for the Custom CNN Model (3-layers) is 62.01% , the train accuracy is only 61.42%. The Custom CNN Model (5-layers) has a train accuracy of 63.75%. The ability to categorize feelings will be useful in gauging a candidate's emotional and mental preparedness for various test situations and questions.

While this study relies on data from a single source, future research may employ webcams or cameras to collect data in real time. After analyzing a candidate's emotional response in a psychometric assessment, it can be compared to the emotional responses of other candidates; this facilitates comparisons between candidates during the hiring process.

Combining this system with Internet of Things techniques will increase its applicability across many sectors. Narcotics officers, for instance, could use such data in real time to analyze and detect the subject's emotional reaction, which could help them identify possible drug use. This will aid the Narcotics Officers in their investigation of any possible drug trafficking or abuse.

References

Borsboom, D., Cramer, A. O. et al. (2013). Network analysis: an integrative approach to the structure of psychopathology, *Annual review of clinical psychology* **9**(1): 91–121.

Breuer, R. and Kimmel, R. (2017). A deep learning perspective on the origin of facial expressions, *arXiv preprint arXiv:1705.01842* .

Connie, T., Al-Shabi, M., Cheah, W. P. and Goh, M. (2017). Facial expression recognition using a hybrid cnn–sift aggregator, *International Workshop on Multi-disciplinary Trends in Artificial Intelligence*, Springer, pp. 139–149.

Cornejo, J. Y. R., Pedrini, H. and Flórez-Revuelta, F. (2015). Facial expression recognition with occlusions based on geometric representation, *Iberoamerican Congress on Pattern Recognition*, Springer, pp. 263–270.

Jain, D. K., Zhang, Z. and Huang, K. (2020). Multi angle optimal pattern-based deep learning for automatic facial expression recognition, *Pattern Recognition Letters* **139**: 157–165.

Jung, H., Lee, S., Yim, J., Park, S. and Kim, J. (2015). Joint fine-tuning in deep neural networks for facial expression recognition, *Proceedings of the IEEE international conference on computer vision*, pp. 2983–2991.

Li, B. Y., Mian, A. S., Liu, W. and Krishna, A. (2013). Using kinect for face recognition under varying poses, expressions, illumination and disguise, *2013 IEEE workshop on applications of computer vision (WACV)*, IEEE, pp. 186–192.

Li, J., Wang, Y., See, J. and Liu, W. (2019). Micro-expression recognition based on 3d flow convolutional neural network, *Pattern Analysis and Applications* **22**(4): 1331–1339.

Li, Y., Zeng, J., Shan, S. and Chen, X. (2018). Occlusion aware facial expression recognition using cnn with attention mechanism, *IEEE Transactions on Image Processing* **28**(5): 2439–2450.

Lopes, A. T., De Aguiar, E., De Souza, A. F. and Oliveira-Santos, T. (2017). Facial expression recognition with convolutional neural networks: coping with few data and the training sample order, *Pattern recognition* **61**: 610–628.

Mahata, J. and Phadikar, A. (2017). Recent advances in human behaviour understanding: A survey, *2017 Devices for Integrated Circuit (DevIC)* pp. 751–755.

Olson-Buchanan, J., Bryan, L. K. and Thompson, L. F. (2013). *Using industrial-organizational psychology for the greater good*, Routledge New York, NY.

Reddy, C. V. R., Reddy, U. S. and Kishore, K. V. K. (2019). Facial emotion recognition using nlpca and svm, *Traitement du Signal* **36**(1): 13–22.

Siddiqi, M. H., Ali, R., Khan, A. M., Kim, E. S., Kim, G. J. and Lee, S. (2015). Facial expression recognition using active contour-based face detection, facial movement-based feature extraction, and non-linear feature selection, *Multimedia Systems* **21**(6): 541–555.

Varma, S., Shinde, M. and Chavan, S. S. (2020). Analysis of pca and lda features for facial expression recognition using svm and hmm classifiers, *Techno-Societal 2018*, Springer, pp. 109–119.

Yu, Z. and Zhang, C. (2015). Image based static facial expression recognition with multiple deep network learning, *Proceedings of the 2015 ACM on international conference on multimodal interaction*, pp. 435–442.

Zhang, K., Huang, Y., Du, Y. and Wang, L. (2017). Facial expression recognition based on deep evolutional spatial-temporal networks, *IEEE Transactions on Image Processing* **26**(9): 4193–4203.