National College of Ireland

Automated Facial Expression Detection using Machine Learning Algorithm

MSc Research Project
Data Analytics

Kalaikamal Mohan
x16145992

School of Computing
National College of Ireland

Supervisor: Mr.Vikas Tomer

National College of Ireland
Project Submission Sheet – 2017/2018
School of Computing

Student Name:	Kalaikamal Mohan
Student ID:	x16145992
Programme:	Data Analytics
Year:	2018
Module:	MSc Research Project
Lecturer:	Mr.Vikas Tomer
Submission Due Date:	15/09/2018
Project Title:	Automated Facial Expression Detection using Machine Learning Algorithm
Word Count:	5974

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 encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	
Date:	3rd September 2018

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
3. 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):	

Contents

1 Introduction **1**
1.1 Motivation . . . 2
1.2 Significance of outcome . . . 2
1.3 Research Question . . . 2
1.4 Research objectives . . . 2

2 Related Work **3**
2.1 Review on Feature Extraction method applied on Face image . . . 3
2.2 Applied Feature reduction technique for facial features . . . 4
2.3 Classification algorithm applied on Facial image features . . . 5

3 Methodology **6**
3.1 Design Specification . . . 7

4 Implementation **7**
4.1 Environmental setup . . . 7
4.2 Pre-processing the image . . . 8
4.3 Feature Extraction stage from the pre-processed image . . . 8
4.4 Applying PCA for feature Reduction . . . 10
4.5 Applied Classification Models on Facial Features . . . 10

5 Evaluation **13**
5.1 Confusion Matrix . . . 13
5.2 Over all Classification Accuracy, Sensitivity and Specificity . . . 13
5.3 F-Score and Kappa Statistics . . . 14
5.4 Discussion of Multi-class SVM classifier Results . . . 15

6 Conclusion and Future Work **16**

Automated Facial Expression Detection using Machine Learning Algorithm

Kalaikamal Mohan
x16145992
MSc Research Project in Data Analytics

3rd September 2018

Abstract

The main aim of the project is to find the novel feature extraction technique along with the suitable machine learning algorithm to detect the facial expression from the face image. This project mainly deals to overcome the drawbacks of existing feature extraction along with machine learning algorithm applied for facial image to detect the emotions. The scope of the project is to do image classification with lesser image by applying machine learning models rather than Deep learning methodologies, whereas the proposed technique can be applied to any face related classification when there is lack of images. In this paper, I have implemented a novel feature extraction method for face image which is the combination of both wavelet and geometrical features to overcome the existing drawbacks of standalone features. Further, those features are trained with different classifiers such as Support Vector Machine (SVM), K-Nearest Neighbour (K-NN) and Single Hidden Layer Feed Forward Neural Network (SHLFNN), and found that Multi-class SVM stands better for the extracted features. The evaluation method has carried out with different metrics such as Accuracy, Sensitivity, Specificity, F-Score and Kappa Statistic to support the proposed methodology.

Keywords: Machine Learning, Deep Learning, KNN, SVM, SHLFNN, PCA

1 Introduction

Machine learning is basically classified into 3 categories such as Regression, Classification and Clustering. Also, all the algorithm will work based on supervised, semi supervised and unsupervised learning methods. In this project, I have chosen to work with supervised classification for the facial image expression detection. According to (SONMEZ and ALBAYRAK; 2016), there are 7 universal emotions are such Happy, Angry, Sad, Fear, Surprise, Neutral and Disgust. The most challenging part is extracting the features from facial image and applying suitable algorithm for that to classify the emotion from facial image. So far, the state of art in image classification using machine learning algorithm is applying wavelet transform on the image and training those features with Artificial Neural Network with specific optimization algorithm, but still it takes longer time to classify emotions.This project brings out novel feature extraction method which overcomes the

drawbacks of existing methods, and the same is followed with machine learning algorithm to stand out better. In the following, we will discuss about the motivation, significance of the project and proposed methodology for detecting the facial expression.

1.1 Motivation

As the technology grows rapidly, the research on human-computer interaction is showing drastic change which leads to greater extent. All the application areas based on human computer interaction has always demand for innovation, and so, I have chosen project where I can contribute to some extent for human face and computer interaction(Bartlett et al.; 2003). Consider yourself that you are using the mobile/laptop and working on it with front camera switched on, so whenever your facial expression changes the computer can read your emotion and respond accordingly. Interestingly, nowadays the Facial expression detection has applied for e-learning systems to detect the emotion of the students and based on that the lecture and presentation can be modified (Ayvaz et al.; 2017). It also applied to robotic and social networking such as Face emotion tagging, automatic caption, etc.

1.2 Significance of outcome

So far, the machine learning and deep learning algorithm has implemented for Facial Expression Detection from the face image. The difference between the machine learning implementation and deep learning implementation for the facial expression detection is about the number of images and its features used to classify the emotions. In deep learning it has advantage of extracting feature by itself, but whereas it requires at least thousand images for each label, and also for implementing the deep learning model we require higher Graphical Processor Unit (GPU) systems (Ongsulee; 2017). Though there are many machine learning algorithms which are used for facial expression detection, each approach has drawbacks with feature sets and selection of classifier, hence after proper research analysis, we have implemented novel feature extraction which is the combination of both wavelet texture features and facial anthropometric features. Also, the proposed model uses less number of images to extract all the possible features from face image and trains the classifier model to achieve better accuracy. Though deep learning methodologies gives very good accuracy for many image classification problem (Lopes et al.; 2017) (Zeng et al.; 2018), it still needs thousands of images for each labels which is not always feasible (Wang and Kosinski; 2017), whereas the proposed methodology has most significant outcome that it can be used on any application of facial image detection with less number of images to train the classifier model. In this way, the proposed methodology will aid to solve many image classification problems.

1.3 Research Question

To what extent the emotion of the person can be identified from combination of both texture and geometric face image features along with suitable machine learning algorithm?

1.4 Research objectives

To accomplish the research to what extent the emotion of the person can be identified through facial image, we have first analysed the state of art methods which has performed

for the facial expression detection. In each methodology we have done research and found few drawbacks on existing standalone feature extraction technique such as wavelet features, edge detected features, texture features and anthropometric features. After careful analysis, the novel feature extraction technique has implemented on facial image to overcome the existing drawbacks and further the extracted features have proceeded with feature engineering such as Principal component Analysis (PCA) to find the important components. At last, the dimension reduced features are trained with classifiers such as K-NN, SVM and SHLFNN. In following, the up-to-date literature review which has done on the facial image expression detection has discussed.

2 Related Work

In recently years, many authors have discussed and proposed many methodologies to analyze the face image, detect the facial parts, extract the features, classify the features to detect the emotions. In each paper, each author has different approach in detecting the facial expression from the face image. In the following, the review has done on how the feature extraction has done from face image, how feature reduction technique helped to reduce the dimension, how and what classification algorithm has applied for the facial features to detect the emotions.

2.1 Review on Feature Extraction method applied on Face image

The feature extraction is the step which starts after loading the images in to the tool, where each tool has different package to work with image and extract features. Many researches, has applied different technique such edge detection, Region of interest detection from face, wavelet feature extraction, texture feature extraction, etc. In the following, the different feature extraction technique applied on the face image has been analyzed.

In paper (Yang et al.; 2018), the author used has done face region detection using HAAR method, through this the eyes and mouth region are identified from the face image. From the detected region, the Sobel edge detection method is applied to detect the facial features. Similarly, the author (Tsai et al.; 2010) (Hong and Xuanbing; 2010) has done another approach, the face is detected first based on the Haar-like features method. Many approaches such as k-level HAAR wavelet decomposition and Stationary Wavelet entropy (SWE) techniques are applied to extract the shape and texture features through the wavelet transform (Hong and Xuanbing; 2010). The Haar-like features is combined to the Self quotient image (SQI) filter, together is applied for the image which helps to identify the face from the image and then the feature is extracted mainly by using Discrete cosine transform (DCT). Interestingly, the author is used Angular Radial Transform (ART) and Gabor Filter (GF) along with DCT to extract the facial features(Tsai et al.; 2010). In this paper, the author mainly concentrated on frown expression detection which are such as anger and sad. The approach involves 3 stages. The first stage is feature extraction which has following steps such as skin color segmentation, Face detection, Eyes Detection, Lip Detection(Ramalingam et al.; 2018). The good skin color segmentation results to face region detection because all the face regions will have some unique color compare to other objects. Once the face is detected, the Region of interest such as eye,

mouth and lips are detected (Ramalingam et al.; 2018) (Verma and Singh; 2011) (Kundu and Saravanan; 2017). The feature extracted in this approach is termed as patch-based Gabor features(Ramalingam et al.; 2018). Once after the ROI is detected using MaZda, which is software package for texture analysis, is applied on the ROI detected region to extract the texture features. This approach is mainly compared with wavelet features and stands better but the time taken and complexity level of feature extraction method is high(Verma and Singh; 2011).

Finding the Region of interest and doing feature extraction is the approach which many authors have followed to classify the emotion from face image. In this paper(Rani and Muneeswaran; 2018), the author proposed a approach for finding the Region of Interest (ROI) in the face. Once after segmentation is done, the author(Kundu and Singh; 2013) has implemented the histogram-based threshold technique and following to that the geometric feature extraction technique is applied on the segmented part of image to get features.

The Viola jones algorithm is the one which is mainly used to detect the facial regions. The author (Rashid et al.; n.d.) has done image segmentation using Viola jones algorithm and then the pre-processing technique is applied based on thresholding and histogram equalization (Rashid et al.; n.d.) (Ali Khan et al.; 2015). Also, the vertical cross correlation is applied to extract the features from the segmented image (Rashid et al.; n.d.). The Haralick features is is also easily extracted and imported (Ali Khan et al.; 2015). In this paper(Kim et al.; 2017), the proposed algorithm for facial expression detection is mainly evaluated with F-1 score which is the harmonic average that calculates for precision and accuracy. In this paper(Kim et al.; 2017). In this paper (Zhang et al.; 2015) it starts from giving input of image and from the input image the features are extracted.Though ROI detected and feature extraction gives more information, the wavelet transform feature gives the significant face image features(Wang et al.; 2018). According to author (Wang et al.; 2018), the approach of feature extraction is mainly to minimized the sensitivity and to increase the robustness of classifier and that can be achieved by hyper parameters adjustment. The stationary wavelet is one among the feature set which works based on wavelet decomposition and from which the features are extracted (Wang et al.; 2018).

Another interesting approach of feature extraction is geometrical features, the author (Lekdioui et al.; 2017), has used the IF algorithm to detect the 49 landmarks from the face image which includes eyes, nose, mouth, eyebrows, etc. Once the landmarks are detected the geometric distances are extracted and used as features, which mainly worked well for images irrespective of color, noise, intensity, etc (Lekdioui et al.; 2017). Also, the anthropometric features are concatenated to get the texture and shape features, whereas, the author (Howard et al.; 2017) have used geometrical distance feature to detect the age classification.

2.2 Applied Feature reduction technique for facial features

The feature reduction technique is mainly applied for the feature sets which are high in dimension. It is mainly applied to avoid the classifier model to meet over fit, because if the model is overfitted then it will have more variation and less biased, so that it can identify the emotions correctly. The mostly applied feature reduction for the image

features is Principle component Analysis (PCA). The PCA. In this paper (Al-Sumaidaee et al.; 2017), features are taken for dimension reduction which was done by PCA (Al-Sumaidaee et al.; 2017). From this, it gives information that PCA applied on texture features works well and can be further used for classification to get better accuracy and less training time.

2.3 Classification algorithm applied on Facial image features

The Artificial Neural Network is the most promising classification algorithm which is applied to many image classification problems (Madani; 2008). In that, exclusively the Single Hidden Layer Feed Forward Neural Network (SHLFNN) with controlling parameters and along with the algorithm specific parameter is used for classifying the wavelet features (Wang et al.; 2018). The ANN model is trained with features which are extracted from the ROI detected parts of image (Kundu and Saravanan; 2017), In this approach, the author has used three hidden layers and achieved accuracy of 83.3% for face emotion classification. Also, author (Yang et al.; 2018) has applied the Neural network with number of hidden layer as one, which has 20 neurons in it. In paper (Kundu and Singh; 2013), the author has used the anthropometric features to train with ANN and achieved accuracy of 75%. These strong evidences are used to decide that ANN is suitable for texture features and geometric features.

Once the Haar-like feature set is ready the Support vector machine (SVM) classifier is used to train and classify the images which gives around 98% accuracy (Tsai et al.; 2010). For the patch based gabor features, the author (Ramalingam et al.; 2018) trained the features with nonlinear neural network which supports with backpropagation algorithm, also the Multi class SVM and Ensemble classifiers are implemented to find the comparison result. The feature extracted through HOG and local binary pattern histogram (LBPH) are combined using mutli-kernel function and trained with the SVM classifier to achieve better accuracy (Zhang et al.; 2015).The features extracted through histogram equalization has trained with multi-class SVM classifier (Ali Khan et al.; 2015) (Al-Sumaidaee et al.; 2017), in this implementation the non-linear kernel function is mainly used with SVM classifier to classify the emotions and gives accuracy of 86.21%(Ali Khan et al.; 2015). As the Multi-class SVM classifier is used for both linear and non-linear feature, it has worked for the HOG features well and yields F1-Score of 0.875 (Kim et al.; 2017). In this paper (Lekdioui et al.; 2017) the texture which are derived by geometric feature are extracted using HOG descriptors, then trained based on multi-class SVM classifier gives classification accuracy of 96%. Through this analysis, it is concluded that the texture and anthropometric features can be classified by multi-class SVM.

The K-Nearest Neighbor is the instance classification algorithm which is used for the image classification using Euclidean distance calculation method. For the features extracted from the histogram equalization is trained using KNN along with vertical cross correlation, this approach gives classification accuracy of 93% (Rashid et al.; n.d.). Interestingly, the author (Verma and Singh; 2011), has used the texture features which is extracted from face image and then compared with two main classification algorithm such as KNN and SVM are used to evaluate the result. From the analysis, it has found that the SVM works better for the texture features (Verma and Singh; 2011). Also, the author (Kumar and Basha; 2014) used the wavelet features to classify using K-NN algorithm and

achieved decent classification accuracy. From these analysis, it gives the information that the K-NN works well for image features.

3 Methodology

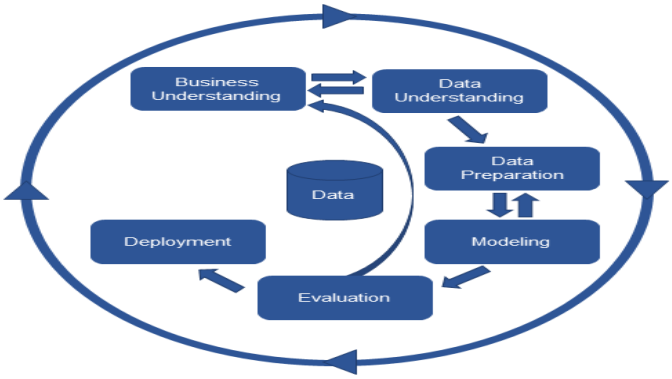

Figure 1: CRISP-DM Methodology

CRISP- DM is the framework and which has been used as methodology for the project implementation. The CRISP-DM has steps such as Business and Data understanding, Data preparation, Modelling, Evaluation and Deployment for Data mining projects (Nadali et al.; 2011). In the following, it was well explained how the project connected to CRISP-DM methodology, what tools were used for project and what innovative approach has applied to bring solution for this research.

In this research, the business understanding relates to our research objective which is detecting the facial expression from the face images. The Data understanding stage starts from acquiring the data and getting knowledge about it. The source data for our research project is face image and it has taken from public access portal which was provided by Ghent university as Karolinska Directed Emotional Faces (KDEF) documents. Each image will show emotions such as Happy, Sad, Fear, Anger, Neutral, Surprise and Disgust. In this dataset, we have around 490 images which has 70 images for each emotion.

In our project, the data preparation stage is involved with feature extraction and feature engineering. The innovative approach of this project has contributed at this stage by extracting the novel features which has the combination of the wavelet and anthropometric features. The stationary wavelet gives the image details such edge detection, vertical details and horizontal details. From this the coefficient of wavelet is extracted and used as feature set (Wang et al.; 2018), (Hong and Xuanbing; 2010). The anthropometric features give the geometrical distances for each emotion (Hernandez-Matamoros et al.; 2016). Once features are ready it has used further for modelling.

The modelling stage of the project comprised of the classification algorithm which has applied to the extracted feature dataset. For this project I have implemented three classification machine learning algorithm which are K-Nearest Neighbour (KNN), Support Vector Machine (SVM) and Single Hidden Layer Feed Forward Neural Network (SHLFNN), the reason for choosing the classifier is mainly based on the research analysis

done so far and explained below. From the literature review analysis, though many classification algorithms have applied on the face image, only few classifiers result out better accuracy for texture and geometric features (Kumar and Basha; 2014), (Lekdioui et al.; 2017) (Kundu and Saravanan; 2017) (Wang et al.; 2018) (Tsai et al.; 2010) (Kundu and Singh; 2013)from the analysis, it has found the wavelet features works well for classifiers such as K-NN, Multi-class SVM and Single Hidden Layer Feed Forward Neural Network (SHLFNN) (Kumar and Basha; 2014) (Lekdioui et al.; 2017) (Kundu and Saravanan; 2017) (Wang et al.; 2018) (Tsai et al.; 2010) (Zhang et al.; 2015) (Kundu and Singh; 2013)Also, there is a proven result that Anthropometric features is also well suited with Multi-Class SVM classifier (Lekdioui et al.; 2017).

The Evaluation stage of our project consists of finding the evaluation metrics for three models such as K-NN, SVM and SHLFNN which were applied to our image features. The metrics for classification generally derived from the confusion matrix, once the confusion matrix is generated the Accuracy, Sensitivity, Specificity, F- Score and Kappa statistic (Ferri et al.; 2009) (Ballabio et al.; 2018).

3.1 Design Specification

The Design of our project is the framework with stages such as pre-processing the input image, extracting the wavelet and anthropometric features, applying PCA for the features, Modelling the features and Evaluating the metrics for the applied models. The below flow chart explains the stages of implementation.

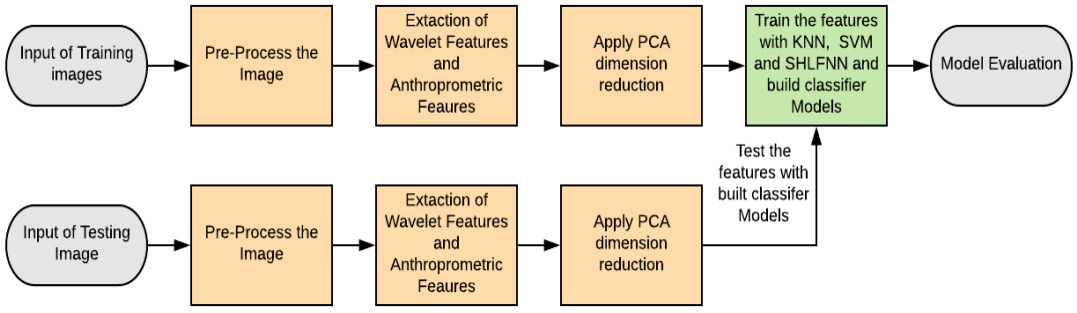

Figure 2: Proposed Work Flow

4 Implementation

The implementation of the project is the most important part of the project which should be done with utmost attention. In this project, the implementation has done as follow with the usage of various tools and algorithms.

4.1 Environmental setup

In this project, different tools were explored for all stages of implementation. Initially, for pre-processing the image we used both MATLAB and python (Reddy et al.; 2017)(Annala

et al.; 2018), then 'R Studio' for PCA and modelling.

4.2 Pre-processing the image

The pre-processing stage is the simple step which was carried on the input of image to resize the dimensions to maintain the consistency through out the dataset.

4.3 Feature Extraction stage from the pre-processed image

The feature extraction is the challenging task in this project, as we need to extract features which is the combination of both wavelet and anthropometric features. As it has mentioned already, the novel approach of feature extraction is to overcome the drawbacks of the existing stand alone feature sets.

Reason for proposing Novel Feature Extraction: Though the wavelet features give all the edge and texture feature of the image, it has some drawbacks for face image which has different skin tone and different age (Wang et al.; 2018) (Howard et al.; 2017). Whereas, the geometric features give all the possible distance measurement between face objects, but it will not hold any edge details of the image. Hence, in this project we have used combination of both wavelet and geometric features. In the following, the feature extraction of both wavelet and anthropometric is explained.

Wavelet Feature extraction: For wavelet feature extraction, the MATLAB has used which has wavedec2 function installed with it. The wavelet feature extraction is done based on image decomposition by passing High pass and Low pass filter on the image and the process repeated to subsequent decomposition level which gives required wavelet co-efficient. The high pass filter is applying high frequency pixel on the image and passing the high pixel information of image through filter, and whereas, the low pass filter is applying low frequency pixel on the image to allow the low pixel information of image to pass through filter (Wang et al.; 2018) (Hong and Xuanbing; 2010). In this project, the level 2 decomposition is applied which gives the required wavelet coefficients. The wavelet transform on the pre-processed image follows the below steps.

Step 1: The pre-processed image is loaded as input.

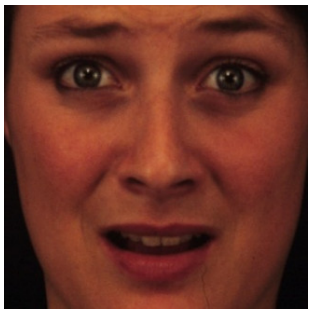

Figure 3: Face image with Fear Emotion

Step 2: The Wavelet decomposition level 1 is applied on the above image, In below, the first Level 1 Decomposition gives the approximation coefficient of level 1 decompos-

ition. Following to that, the horizontal, vertical and Diagonal detail of the coefficient is extracted.

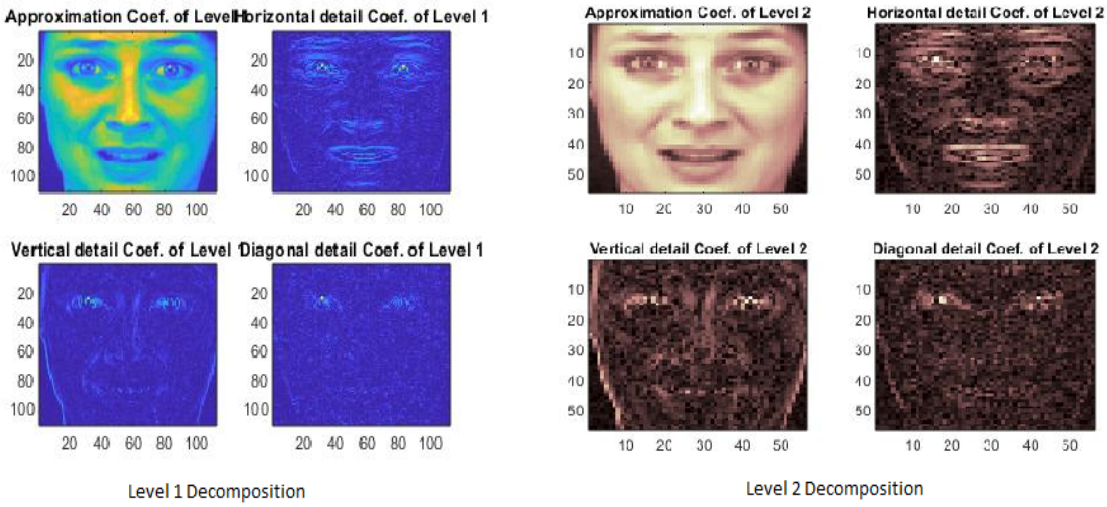

Figure 4: Level 1 and Level 2 Decomposition of image

Step 3: Further, the level 2 decomposition is applied subsequently to get wavelet information in deeper. In the above figure, the level 2 decomposition of image is displayed with horizontal, vertical and diagonal detail of image is shown, from which the wavelet coefficient is extracted.

Step 4: From the sub band of images, the wavelet features is extracted and loaded as Excel file.

Anthropometric Feature extraction: The face anthropometric feature is the measurement of geometrical distance between the face objects. For this approach, the Python code is used to implement the feature extraction process. In python, the face recognition package is used to identify the face and facial landmarks (Annala et al.; 2018). Once the facial land mark is identified, the important geometrical distances such as width of left and Right Eye, length of left and right eye, distances between eye and eyebrow, length and width of mouth, slope of eyebrows (Lekdioui et al.; 2017) (Howard et al.; 2017) (Kundu and Singh; 2013). In the following, the step by step process of extracting the geometrical distances is explained

Step 1: The pre-processed image is loaded as input.

Step 2: Using Python code, the face is recognized and the facial landmarks are identified and plotted. In the below Figure, the image with facial landmarks detected, line drawn on landmarks has displayed.

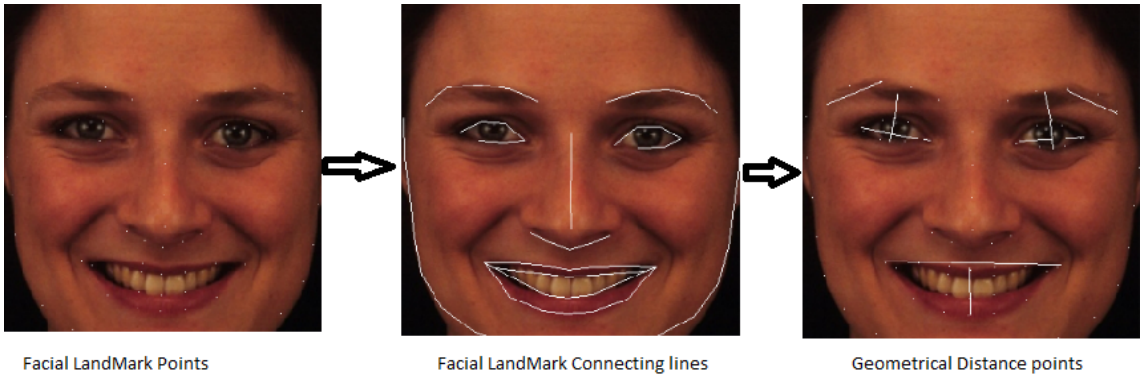

Figure 5: Anthropometric Feature Extraction Stages

Step3: In the step3, as the Fig above, the geometrical distance which are measured is plotted for reference.

Step 4: The geometrical distance is considered as face anthropometric feature sets and it has exported to Excel sheet for further process.

4.4 Applying PCA for feature Reduction

Once the feature sets are ready, Principle components analysis (PCA) is applied using R tool for the feature sets. The main reason for applying dimension reduction technique is that the wavelet features extracted from the image are in higher dimension, and which causes curse of dimension. Though many dimension reduction exists, the PCA applied to image features has given better result in classification task (Al-Sumaidaee et al.; 2017) , and so the PCA has chosen. In the following, how the processed feature sets were used for different classification algorithm and the implementation of the same were discussed.

4.5 Applied Classification Models on Facial Features

Despite of many machine learning classification algorithm exists for image features, in this project the image features are trained and classified with classifier model such as KNN, SVM and SHLFNN. The hold out validation is same for all the models applied in this project, as it takes 80% of data as Train Data and 20% of data as test data.

K-Nearest Neighbour (KNN) Model implementation on Image Features: The K-Nearest neighbour is the instance-based classification algorithm which can be used for both regression and classification problems. The KNN has mainly 3 aspects which are as follows, the first is the output of KNN is very easy to interpret, the second is the calculation time for classifying the test image, and the last is predictive or classify power (Tsoumakas and Katakis; 2007).

The reason behind choosing KNN classification algorithm is, that it works for texture and spatial features, also it has proven result that the KNN gives better accuracy for wavelet features (Kumar and Basha; 2014) (Rashid et al.; n.d.). Also, the KNN captures even the small changes in the features which causes the result, in this project as the geometrical distances are smaller in variation compared to each other, the KNN is well used to work on the features to classify the emotions (Kumar and Basha; 2014). The

KNN model is built based on the following steps using R studio which is explained as Pseudo code.

Step 1: Input the features dataset.

Step 2: The k value is initialized which is the significant value for KNN classifier.

Step 3: The feature sets is arranged in n-dimensional space and new test feature is passed through it.

Step 4: The distances between the test features and the training features is calculated based on distance formula. The distance formula available for KNN is Euclidean, Cosine, Chebyshev, etc. In this project, the Euclidean distance formula is used to find the distances.

Step 5: The number of nearest distances is calculated based on the k values. Once the distances are found the most frequent emotion is assigned to test image. Below image shows accuracy for different 'k' values

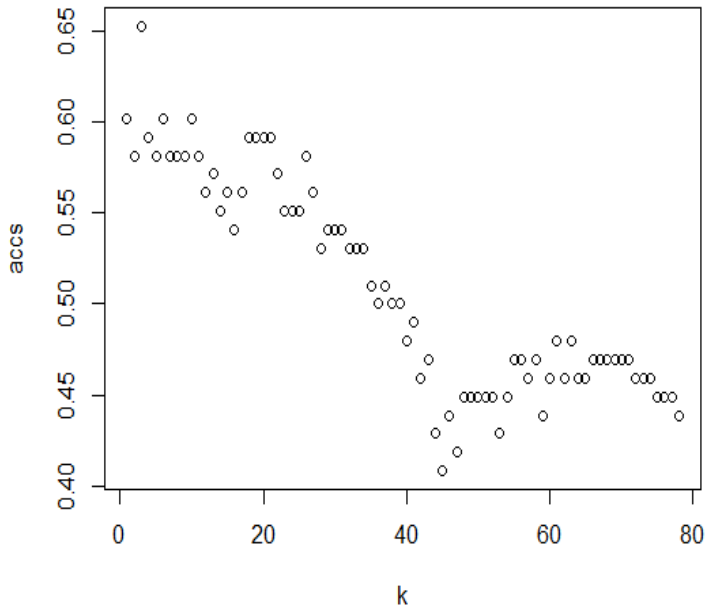

Figure 6: K Value vs Accuracy

Support Vector Machine (SVM) Model implementation on Facial features: The Support vector Machine (SVM) is one among the best suitable classifier for the image features. The SVM classifier basically inspired from the statistical learning methodologies which used statistical technique to classify the target(Tsoumakas and Katakis; 2007).

The reason behind choosing the SVM classifier is that it has shown proven result of good classification result for texture features (Lekdioui et al.; 2017) (Al-Sumaidaee

et al.; 2017). Also, as it uses the statistical technique the geometrical distance features are well classified using the SVM classifier. The SVM classifier basically do the binary classification, in this project the multi-class SVM classifier is used to classify the emotions (Lekdioui et al.; 2017). The pseudo code for the SVM classifier implemented for this project is explained below.

Step 1: The feature set is given as input and it has plotted in n-dimensional space.

Step 2: The hyper plane is created between class features, which is the separable line that has the maximum distance from the hyperplane to either side of the feature class.

Step3: The hyper plane can be created using function such as linear, radial, etc. In this project as our features are linearly separable the hyperplane is created using linear function.

Step 4: Once the model is trained by creating hyper planes, the new test feature set is identified on which class of hyper plane it falls, and the corresponding the emotion is assigned to the test data.

Single Hidden Layer Feed Forward Neural Network implementation on Facial features: The Artificial neural network is the activity which is similar to the function of our human brain. It starts with input layer which takes the image features and works with hidden layer to hit the emotion target. The neural network is named based on the number of hidden layer used, in this project the number of hidden layer used is one and so it is defined as Single Hidden Layer Feed Forward Neural Network.

The reason for choosing Artificial neural network (ANN) in this project is that ANN works well for both linear and non-linear behaviour of the image features, also it has shown proven result for many image classification problems (Wang et al.; 2018), (Madani; 2008) (Kundu and Saravanan; 2017). Specifically, the facial image features such as wavelet features are classified well based on the proper training with ANN (Yang et al.; 2018). In the following, the pseudo code for implementation of ANN is explained below.

Step 1: Input the combined feature set in to the input layer.

Step 2: The number of input layer is equals to number of image features it feeds to network.

Step 3: The input features is multiplied with input weights and hits the hidden layer.

Step 4: The intermediate values from the hidden layer is then multiplied with output weights and hits the target which is the emotion of the image.

Step 5: If the emotion is not wrongly classified, the error function is calculated and the weights are updated to repeat the steps from 2 to 4 till it achieved the best result.

5 Evaluation

Once after the classifier model is built, the performance evaluation is carried out for the models to find how well the emotions are classified from the image features (Ferri et al.; 2009). In this project, the simple split approach is applied which is termed as Hold out method. Based on the hold out method, 80% of the image from the database are used for training and 20% for testing. The evaluation metrics such as confusion Matrix, Classification Accuracy, Sensitivity, specificity, F-Score, ROC Curve and Kappa statistics are measured (Ferri et al.; 2009) (Ballabio et al.; 2018).

5.1 Confusion Matrix

The confusion matrix is base metrics which is used for classification problems. The reason behind choosing the confusion Matrix as one of the metrics is that it gives the classification of each labels, from which the metrics such as Precision, Accuracy, Specificity, Recall and F-score are generated. The values in the confusion matrix is stored as True positive, True Negative, False Positive and False Negative(Ferri et al.; 2009). The confusion matrix is generated for all the classifier models such as K-NN, SVM and SHLFNN which was tested for the test image features.

5.2 Over all Classification Accuracy, Sensitivity and Specificity

The overall classification Accuracy is the measurement which shows the model effectiveness to the test images. It is mainly calculated based on the true classification and false classification rate (Ferri et al.; 2009). The True classification rate is the amount to which the classifier model classified emotions correctly. Conversely, the False classification rate is the amount to which the classifier model classified the emotions incorrectly (Ballabio et al.; 2018).

The sensitivity is the measurement of True positive rate which tells how frequently the classifier model classifies the emotions correctly. In contrast, the specificity is about how frequently the model classifies the false emotions correctly, and it is mainly based on False Negative rate (Ferri et al.; 2009). The reason behind choosing the sensitivity and specificity is that if the happy face classified as Sad face, which is the extreme points, then it will have more negative impact in real time application. Hence, to ensure the proposed model is well used in the real time application, along with Classification accuracy, the Sensitivity and Specificity are calculated for the built models such as K-NN, SVM and SHLFNN (Ballabio et al.; 2018).

	K-Nearest Neighbour (KNN) (%)	Single Hidden Layer Feed Forward Neural Network	Support Vector Machine (SVM)(%)
Over All Accuracy	62.24	65.31	**85.71**
Over All Sensitivity	61.71	64.71	85.28
Over All Specificity	93.14	93.85	97.28

Figure 7: Overall Accuracy, Sensitivity and Specificity

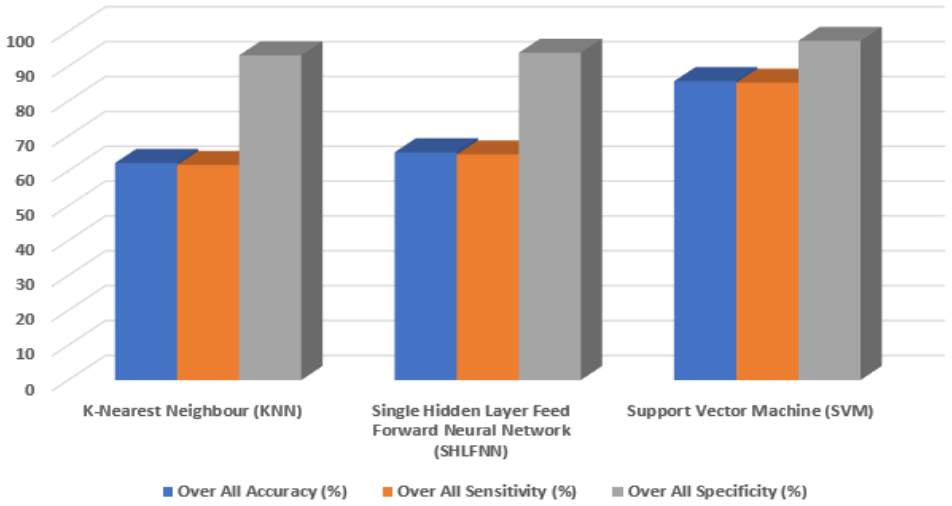

Figure 8: Comparison chart of overall Accuracy, Sensitivity and Specificity

In the above table, the overall accuracy, specificity and sensitivity across different model implemented on the features have displayed. Form the table, it shows that K-NN is the least workable for the extracted image feature sets because the combination of features becomes sensitive to KNN Model. Following to that Single Hidden Layer Feed Forward Neural Network (SHLFNN) is better compared to KNN in terms of classification accuracy, specificity and Sensitivity. It is found that the Multi-Class Support Vector Machine (SVM) stands out better when compared to other classifiers built for our image features.

5.3 F-Score and Kappa Statistics

The F-score is the measure which is calculated based on the harmonic average of precision and recall. The F-score ranges from 0 to 1, where the value of 1 is the best score and the 0 is the worst score(Ferri et al.; 2009) (Ballabio et al.; 2018). The kappa statistics is the measurement which shows how good the agreement between test and train data in terms of classification(Ferri et al.; 2009) (Ballabio et al.; 2018). In the following, the F-score and Kappa statistics value of the classifier models applied on the image features is given.

	K-Nearest Neighbour (KNN)	Single Hidden Layer Feed Forward Neural Network (SHLFNN)	Support Vector Machine (SVM)
F-Score	0.61	0.65	0.85
Kappa statistic	0.57	0.59	0.83

Figure 9: F-Score and Kappa Statistic

From the result, it is clear the multi-class SVM which is with type of C-Classification used for the features set has worked well when compared to other classifiers. In addition, the multi-class SVM shows the classification rate for each emotion higher than other classifiers. In the following, the confusion matrix and the individual sensitivity and accuracy of Multi-class SVM classifier is given.

5.4 Discussion of Multi-class SVM classifier Results

The confusion Matrix and the statistical results such as Specificity, Sensitivity of SVM classifier are discussed

Confusion Matrix of Multi-class SVM classifier

Confusion Matrix and Statistics

	Reference						
Prediction	Angry	Disgusted	Fearful	Happy	Neutral	Sad	Surprised
Angry	12	3	0	0	0	0	0
Disgusted	0	10	0	1	0	2	0
Fearful	0	0	14	0	0	0	0
Happy	0	0	0	12	0	1	0
Neutral	2	1	0	1	14	2	1
Sad	0	0	0	0	0	9	0
Surprised	0	0	0	0	0	0	13

Figure 10: Confusion Matrix of Multi-Class SVM classifier

```
          Accuracy : 0.8571
            95% CI : (0.7719, 0.9196)
   No Information Rate : 0.1429
   P-Value [Acc > NIR] : < 2.2e-16

             Kappa : 0.8333
 Mcnemar's Test P-Value : NA

Statistics by Class:

                     Class: Angry Class: Disgusted Class: Fearful Class: Happy Class: Neutral Class: Sad
Sensitivity                0.8571           0.7143         1.0000       0.8571         1.0000    0.64286
Specificity                0.9643           0.9643         1.0000       0.9881         0.9167    1.00000
Pos Pred Value             0.8000           0.7692         1.0000       0.9231         0.6667    1.00000
Neg Pred Value             0.9759           0.9529         1.0000       0.9765         1.0000    0.94382
Prevalence                 0.1429           0.1429         0.1429       0.1429         0.1429    0.14286
Detection Rate             0.1224           0.1020         0.1429       0.1224         0.1429    0.09184
Detection Prevalence       0.1531           0.1327         0.1429       0.1327         0.2143    0.09184
Balanced Accuracy          0.9107           0.8393         1.0000       0.9226         0.9583    0.82143
                     Class: Surprised
Sensitivity                    0.9286
Specificity                    1.0000
Pos Pred Value                 1.0000
Neg Pred Value                 0.9882
Prevalence                     0.1429
Detection Rate                 0.1327
Detection Prevalence           0.1327
Balanced Accuracy              0.9643
```

Figure 11: Overall Statistics of Multi-Class SVM Classifier

It can be interpreted that that SVM classifier has classified the emotions with good classification rate. To highlight the results of model built for our features, the emotions such as Fearful and Neutral has classified to 100% accuracy. When compared to other algorithm, it gives the result which is outstanding.

The sensitivity and specificity for all the emotions has shown very accuracy with Multi-class classification result. It is also clear that the sensitivity and Specificity of Fearful emotions has been in successful rate, which means that proportions of emotions which are to be classified as Fearful has correctly classified, and similarly, the rest of the emotions also not classified wrongly as Fearful. This shows the very good result of using Multi-class SVM classifier for our project.

6 Conclusion and Future Work

The project result shows that the novel feature extraction from face image has been trained and classified with 3 machine learning algorithms and found that extracted features work better for Multi-Class SVM. Also, as stated before the novel feature extraction which is the combination of both wavelet and anthropometric features has overcome the drawbacks of existing standalone features for certain emotions. The proposed novel feature extraction method is implemented and those features were trained with KNN, SHLFNN and SVM classifiers. To conclude from the result, the combination of wavelet and anthropometric features along with multi-class SVM classifier has given good very good accuracy of 85.71%. Though, the proposed methodology has not met the overall accuracy of existing state of art methods, it is noted that the proposed algorithm classifies the Fearful and Neutral emotion with 100% accuracy which exceeds the state of art methods. Also, the specificity and sensitivity of classifying the Fearful emotions from the face image is 100%. It means, that the Multi-class SVM classifier built for the proposed features never miss-classified the Fearful emotions, also never classified other emotions as Fearful emotion. The future work can be done analyzing the emotions which are miss-classified and trying with adding few more features following with these standard classifiers. Also, the work can be further taken up by trying the proposed algorithm with different image database which has images that shows large variation in quality of image, pose of the emotion, etc.

Acknowledgement

I sincerely thank my supervisor Vikas Tomer for supporting me in all aspects to complete my project. Also, he has encouraged my project to publish in upcoming conference in future and I wish to do the same. Also, I wish to acknowledge the data source which was posted by Ghent University for public access.

References

Al-Sumaidaee, S., Abdullah, M., Al-Nima, R., Dlay, S. and Chambers, J. (2017). Multi-gradient features and elongated quinary pattern encoding for image-based facial expression recognition, *Pattern Recognition* **71**: 249–263.

Ali Khan, S., Riaz, N., Akram, S. and Latif, S. (2015). Facial expression recognition using computationally intelligent techniques, *Journal of Intelligent & Fuzzy Systems* **28**(6): 2881–2887.

Annala, L., Eskelinen, M., Hämäläinen, J., Riihinen, A. and Pölönen, I. (2018). Practical approach for hyperspectral image processing in python, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 42, International Society for Photogrammetry and Remote Sensing.

Ayvaz, U., G¨ur¨uler, H. and Devrim, M. O. (2017). Use of facial emotion recognition in e-learning systems, *Information Technologies and Learning Tools* **60**(4): 95–104.

Ballabio, D., Grisoni, F. and Todeschini, R. (2018). Multivariate comparison of classification performance measures, *Chemometrics and Intelligent Laboratory Systems* **174**: 33–44.

Bartlett, M. S., Littlewort, G., Fasel, I. and Movellan, J. R. (2003). Real time face detection and facial expression recognition: Development and applications to human computer interaction., *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on*, Vol. 5, IEEE, pp. 53–53.

Ferri, C., Hernández-Orallo, J. and Modroiu, R. (2009). An experimental comparison of performance measures for classification, *Pattern Recognition Letters* **30**(1): 27–38.

Hernandez-Matamoros, A., Bonarini, A., Escamilla-Hernandez, E., Nakano-Miyatake, M. and Perez-Meana, H. (2016). Facial expression recognition with automatic segmentation of face regions using a fuzzy based classification approach, *Knowledge-Based Systems* **110**: 1–14.

Hong, Z. and Xuanbing, Z. (2010). Texture feature extraction based on wavelet transform, *Computer Application and System Modeling (ICCASM), 2010 International Conference on*, Vol. 14, IEEE, pp. V14–146.

Howard, A., Zhang, C. and Horvitz, E. (2017). Addressing bias in machine learning algorithms: A pilot study on emotion recognition for intelligent systems, *Advanced Robotics and its Social Impacts (ARSO), 2017 IEEE Workshop on*, IEEE, pp. 1–7.

Kim, S., An, G. H. and Kang, S.-J. (2017). Facial expression recognition system using machine learning, *SoC Design Conference (ISOCC), 2017 International*, IEEE, pp. 266–267.

Kumar, V. and Basha, A. S. A. (2014). Facial expression recognition using wavelet and k-nearest neighbour, *Current Trends in Engineering and Technology (ICCTET), 2014 2nd International Conference on*, IEEE, pp. 48–52.

Kundu, P. and Singh, R. K. (2013). Geometric feature based recognition of facial expressions using ann, *Signal Processing, Computing and Control (ISPCC), 2013 IEEE International Conference on*, IEEE, pp. 1–6.

Kundu, T. and Saravanan, C. (2017). Advancements and recent trends in emotion recognition using facial image analysis and machine learning models, *Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2017 International Conference on*, IEEE, pp. 1–6.

Lekdioui, K., Messoussi, R., Ruichek, Y., Chaabi, Y. and Touahni, R. (2017). Facial decomposition for expression recognition using texture/shape descriptors and svm classifier, *Signal Processing: Image Communication* **58**: 300–312.

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.

Madani, K. (2008). Artificial neural networks based image processing & pattern recognition: From concepts to real-world applications, *Image Processing Theory, Tools and Applications, 2008. IPTA 2008. First Workshops on*, IEEE, pp. 1–9.

Nadali, A., Kakhky, E. N. and Nosratabadi, H. E. (2011). Evaluating the success level of data mining projects based on crisp-dm methodology by a fuzzy expert system, *Electronics Computer Technology (ICECT), 2011 3rd International Conference on*, Vol. 6, IEEE, pp. 161–165.

Ongsulee, P. (2017). Artificial intelligence, machine learning and deep learning, *ICT and Knowledge Engineering (ICT&KE), 2017 15th International Conference on*, IEEE, pp. 1–6.

Ramalingam, V., Pandian, A. and Jayakumar, L. (2018). Facial emotion recognition system–a machine learning approach, *Journal of Physics: Conference Series*, Vol. 1000, IOP Publishing, p. 012028.

Rani, P. I. and Muneeswaran, K. (2018). Emotion recognition based on facial components, *Sādhanā* **43**(3): 48.

Rashid, M., Hasan, M., Yeasmin, N., Shahnaz, C., Fattah, S., Zhu, W. and Ahmed, M. (n.d.). Emotion recognition based on vertical cross correlation sequence of facial expression images.

Reddy, R., Kompala, N. and Ravi, N. (2017). Digital image processing through parallel computing in single-core and multi-core systems using matlab, *Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2017 2nd IEEE International Conference on*, IEEE, pp. 462–465.

SONMEZ, E. and ALBAYRAK, S. (2016). A facial component-based system for emotion classification, *Turkish Journal of Electrical Engineering & Computer Sciences* **24**(3): 1663–1673.

Tsai, H.-H., Lai, Y.-S. and Zhang, Y.-C. (2010). Using svm to design facial expression recognition for shape and texture features, *Machine Learning and Cybernetics (ICMLC), 2010 International Conference on*, Vol. 5, IEEE, pp. 2697–2704.

Tsoumakas, G. and Katakis, I. (2007). Multi-label classification: An overview, *International Journal of Data Warehousing and Mining (IJDWM)* **3**(3): 1–13.

Verma, G. K. and Singh, B. K. (2011). Emotion recognition based on texture analysis of facial expression, *Image Information Processing (ICIIP), 2011 International Conference on*, IEEE, pp. 1–6.

Wang, S.-H., Phillips, P., Dong, Z.-C. and Zhang, Y.-D. (2018). Intelligent facial emotion recognition based on stationary wavelet entropy and jaya algorithm, *Neurocomputing* **272**: 668–676.

Wang, Y. and Kosinski, M. (2017). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.

Yang, D., Alsadoon, A., Prasad, P., Singh, A. and Elchouemi, A. (2018). An emotion recognition model based on facial recognition in virtual learning environment, *Procedia Computer Science* **125**: 2–10.

Zeng, N., Zhang, H., Song, B., Liu, W., Li, Y. and Dobaie, A. M. (2018). Facial expression recognition via learning deep sparse autoencoders, *Neurocomputing* **273**: 643–649.

Zhang, X., Mahoor, M. H. and Mavadati, S. M. (2015). Facial expression recognition using lp-norm mkl multiclass-svm, *Machine Vision and Applications* **26**(4): 467–483.