

Facial Expression Classification using KNN and Decision Tree Classifiers

M Murugappan*¹ A M Mutawa² Sai Sruthi¹ Aya Hassouneh³ Ali Abdulsalam¹ Jerritta S⁴ Ranjana R⁴¹Department of Electronics and Communication Engineering, Intelligent Signal Processing Research Lab,

Kuwait College of Science and Technology (A Private University), Doha, Kuwait

²Computer Engineering Department, College of Engineering and Petroleum, Kuwait University, Kuwait³Kuwait Technical College (K-Tech), Kuwait⁴Department of ECE, Vels Institute of Science and Technology and Advanced Studies (VISTAS), Tamilnadu, India.

Abstract— In recent years, facial emotional expression recognition attracts several researchers for developing intelligent human-machine interface (HMI) system. This present work classify six different facial expressions (happiness, sadness, anger, fear, disgust, and surprise) using two classifiers namely, K Nearest Neighbor (KNN) and Decision Tree (DT) classifiers. Fifty-five undergraduate university students (35 male and 20 female) with a mean age of 23.9 years voluntarily participated in the experiment to acquire six facial emotional expressions using ten virtual markers called Facial Action Units (FAUs). Firstly, Haar-like features are used for detecting the face and eyes in a video-frame using Viola-Jones adaboost classification method. These FAUs are placed on specific location on the subject's face based on facial action coding system (FACS) using a mathematical model. Lucas-Kande optical flow algorithm is used to continuously track the markers positions. Here, the distance between the FAU at the center of the subject face to other markers are calculated and used as a feature for facial expression classification. The one-way analysis of variance with a significance level of $p < 0.01$ is used to validate the extracted features and fivefold cross-validation is performed. Finally, these cross-validated features are used to map six different facial emotional expression using KNN and DT classifiers. In KNN, four distance measures are used to compare the performance of KNN classifier in emotion classification. The mean emotion classification accuracy of 98.03%, 97.21%, is achieved using KNN and DT, respectively. This accuracy of emotional expression detection using an optical flow algorithm give way for designing real-time systems for variety of applications.

Keywords—Face emotion recognition, Virtual Markers, K Nearest Neighbor, Decision Tree

I. INTRODUCTION

Emotions are the key elements in establishing a communication channel between humans and humans and machines. In recent years, machine learning and artificial intelligence method utilize the field of emotion recognition in the design of human-machine interface devices and intelligent robotics. In the literature, different approaches based on physiological signals, speech, gestures, facial expressions are mostly utilized in emotion classification. Conventionally, human emotions can be recognized by many modalities such as biosignals, facial expressions, gestures, and speech. Among these methods, facial expressions are the most common and widely used in several applications, which include mobile phones, animations, biometric systems, affective computing, clinical diagnosis systems, forensic, etc. [1-5, 15-17]. The primary reason behind the success of facial expression recognition is due to the less complicated

analysis, cost-effective, and simple system design requirement compared to other modalities.

Facial expression recognition is more prevalent over several decades, and several international standard databases developed for assisting the researchers in benchmarking their methods on facial expression recognition. Some of the most common types of a database used by researchers are Japanese Female Facial Expressions (JAFPE), Extended Cohn-Kanade database (CK+), Cohn-Kanade database (CK), MMI, Multimedia Understanding Group (MUG), Indian Spontaneous Expression Database, Facial Expression Research Database (FERG), etc. Most of these databases utilize static images/pictures of different emotional expressions and could not comply with our requirement on using virtual markers based facial emotional expression system. Hence, a novel and international standard database are required for this present work. The local region and Webber local descriptor features are used for classifying six basic emotions using three facial image databases (CK+, MMI, and Static Face in the Wild (SFEW), and achieved a maximum recognition rate of 98.62%, 95.58%, and 50%, respectively [5]. The generalized supervised learning model is utilized on two international standard databases (CK+ and MMI) to classify the facial emotional expressions using the Extreme Learning Machine (ELM) classifier. They achieved a maximum emotion recognition rate of 82.44%, 97.72% in CK+, and MMI databases, respectively [6]. A complete review of emotion recognition databases can found the detail in [7].

Researchers have developed several algorithms for emotion recognition based on static facial images, facial images with fixed reflective markers, and virtual markers. A systematic review of different methodologies in facial emotion recognition using facial images can be found in [8]. Some of the most common issues in facial image method is, computationally expensive – since the algorithm process all the pixels of facial image for emotion recognition, and poor performance – the external factors such as lighting, background changes, significantly affects the performance on emotion classification. Some of the most common regions in a face such as lips, eyes, and cheeks are mostly considered for geometric feature extraction in facial emotional expression analysis. These geometric features-based approaches are computationally efficient (required lesser computation time and memory) and reliable compared to static facial images based facial emotion classification. Thereby, most of the research are focused on developing different types of geometric feature extraction algorithms for facial emotion classification.

To overcome these limitations, researchers started using manual markers –reflecting stickers placed on specific locations on subject’s face to identify the emotions through marker movements. However, these methods are not reliable and do not produce good accuracy in emotion detection. The main issues of this method are, subjects feel uncomfortable in placing the markers, and movement of subjects affects marker placement location, and these lead to achieving a lower recognition rate. Virtual markers attract most of the researchers over the recent years, and these markers are placed on a specific location on the subjects face on some type of algorithms, and the marker movements are continuously tracked by the system to detect the corresponding emotion.

Facial expression recognition studies usually involved FACS for facial emotional expressions. According to FACS, there are 40 muscle locations in the human face have been identified as fiducial points for identifying six basic emotional emotions such as happiness, anger, fear, sadness, surprise, and disgust. These fiducial points can be used as markers and called action units (AUs). Thereby, this method provides a way for developing intelligent real-time facial emotion recognition system based on AUs [11]. The number of action units used to detect the facial expression varies on different research works in the literature. A minimum of 4 AUs to a maximum of 512 AUs are used in the literature.

The facial emotion expression recognition algorithms proposed in the literature is mostly focusses on “offline” applications and very few research works have been reported about real-time facial emotion classification algorithms. In a real-time application, the optical flow algorithm is most successful in detecting the facial expressions under uneven lighting, facial rotation or movement, and changes in the background. Because, optical flow algorithm-based method works on the principle of marker information (x,y) rather than all image pixels. Thereby, the time required for detecting different facial emotional expression using optical flow algorithms is less compared to conventional facial images-based emotion recognition methods. In [2], six basic emotions are detected using Lucas-Kande Optical Flow algorithm and achieved a maximum emotion recognition rate of 83.33%.

The applications of different types of machine learning, and deep learning algorithms perform gained a major attention over recent years in emotional expression recognition. Artificial Neural Networks are used for classifying six basic emotions of the subjects using eleven physical markers achieved a maximum mean emotional expression recognition rate of 96% in [2]. Researchers have achieved a maximum mean emotion expression recognition rate of 69.17% using facial image features and Convolutional Neural Network (CNN) in classifying four basic emotions [4]. Besides, Long Short Term Memory (LSTM) based deep learning network is used to classify the facial expressions using conventional features [6, 9]. Recently, a maximum recognition rate of 99.73 is achieved for classifying six emotions using Spatio-temporal features in Long Short Term Memory (LSTM) networks [9].

The main aim of this present work is to design and develop an intelligent real-time facial emotion expression recognition system to classify six expressions using simple statistical features and with ten virtual markers. Thereby, the computation requirement of a system for virtual marker

placement, feature calculation, and decision making could be significantly reduced compared to conventional methods. Besides, this work also aims to retain the marker positions of neutral state to a maximum of $\pm 25^\circ$ rotation with different backgrounds and skin colors. Lastly, the distance features computed for different emotional expressions are used to map to the corresponding emotions using two non-linear classifiers such as K Nearest Neighbor (KNN) and Decision Tress (DT).

II. METHODS AND MATERIALS

The complete research methodology of the present work is shown in Fig 1.

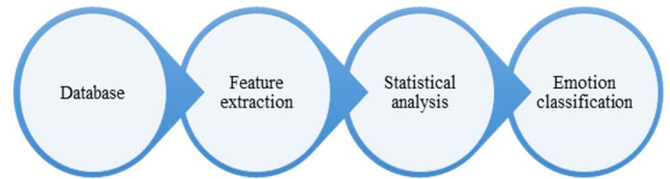


Fig. 1. Flow of facial emotional expression classification

A. Database

To develop an intelligent real-time system for facial emotional expression detection, a quality database is highly essential. In this study, a facial expressions of six emotions are captured through Apple Macbook Pro HD Webcamera with 30 fps. In total, 55 subjects (35 men; 20 women) with a mean age of 22.9 years are voluntarily participated in the experiment. All the experiments have been performed in a controlled environment with a constant room temperature (24°C), room lighting intensity (50 Lux), and with a constant distance between the subject’s face to the camera. The complete flow of data acquisition protocol is shown in Fig 2. A computerized PowerPoint show is developed to guide the subjects in an autonomous way to express their emotional expressions. International affective picture system (IAPS) is used to induce the subjects, and each subject is requested to express each emotion over ten trials, and each trial lasts for the 6-sec duration. A cessation of 10 sec in between the emotional expression is given to the subjects to feel calm and to prepare for the next emotional expression. Thereby, the total time required for one subject to complete all six emotions is seven min. All the subjects are undergraduate students and healthy without having any issues related to muscular or neurological disorders in perceiving or expressing facial emotions.

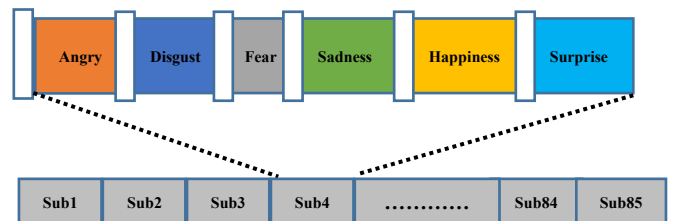


Fig. 2. Data acquisition protocol for emotion recognition

B. Facial virtual marker placement and feature extraction

Several approaches have been proposed by different researchers on face detection. One of the simple methods to

face detection in real-time is proposed by Viola-Jones using Haar-like features. In this study, the subject face is captured by using an HD camera and converted into a greyscale image. This conversion process applied to reduce the computational complexity in processing facial images in facial expression recognition. Then, the greyscale images are used to detect the subject's eye, and ten virtual markers are placed on the subject's face on defined locations using a mathematical model which is proposed by the authors in their earlier work. The complete methodology of the marker placement algorithm can be found in detail in [10]. These ten virtual markers are called FAUs in the facial action coding system. Finally, each marker position (coordinate information) data is transferred to Lucas-Kande optical flow algorithm to trail the changes in marker position during subject emotional expression. Euclidean distance measure is used to calculate the distance between the center marker in the face to each marker (Eqn 1). The Euclidean distance measure is simpler and performs well in distance computation compared to other methods. The computing system automatically stores the marker distances in CSV format during the data acquisition process for further processing. The complete algorithm of face detection, RGB to grayscale conversion, and implementation of an optical flow algorithm is performed through the Open Computer Vision (Open CV) library [12].

$$\text{Distance Measure} = [(X_o - X_v)^2 + (Y_o - Y_v)^2]^{1/2} \text{ -----(1)}$$

where, (X_o, Y_o) is the coordinate of the center marker, and (X_v, Y_v) is the coordinate of a virtual marker.

C. Statistical analysis and cross-fold validation

The distance features of different emotional expressions are statistically tested with a significance level of $p < 0.01$ using a one-way analysis of variance (ANOVA). As a result of the investigation, the distance feature achieved a value of $p = 12.35 \times 10^{-13}$ and a critical value of 4970.239 (F value). Finally, the features are cross-validated using five-fold cross-validation method. Here, four folds data is used for training and the last remaining fold data is used for testing. This procedure will be repeated for five-times to ensure all the training data become a test data in the emotion classification tasks. A detail results of experimental results of statistical analysis of features, and classification of emotional expressions using non-linear classifiers are given in detail in Section III.

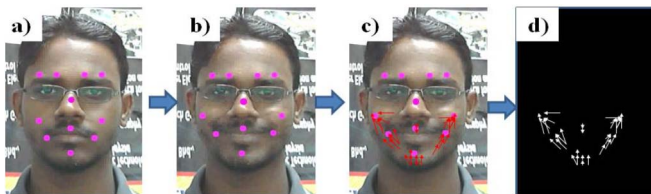


Fig. 3. A sample facial emotional expression

D. K Nearest Neighbor classifier

KNN is a simple nonlinear classifier which is mostly used in several applications, which includes epilepsy detection, driver drowsiness detection, emotion recognition, seizure detection, and many other problems. KNN is basically a non-probabilistic learning algorithm that is used to classify unknown test data based on most similar data among the k-nearest neighbors that are closest to test/unknown data. Different distance measures can be used to measure the distance between the unknown data and each of the trained data such as Manhattan, Euclidean, Minkowski, and Chebyshev. The ability of classification in KNN is mainly dependent on the type of distance measure used [13]. A more detail review of the effect of different measures in KNN based classification can be found in [13]. In this study, the above four different measures are used for classifying the facial emotional expressions, and the mean accuracy of each measure is reported in Section III. Besides the distance measure, the value of k-nearest neighbor has been varied from 1 and 10 with unit step increment, and the value of k-neighbor, which gives the highest accuracy, is listed in the results and discussion section.

E. Decision Tree classifier

Decision Tree (DT) is a supervised machine learning algorithm, and it principally works on the concept of statistical prediction and modeling. This classifier can understand the definitive decision making knowledge from the training data. This classifier is mostly used in data modeling, data mining, machine learning, and decision-making applications. Three performance measures are used to evaluate the classifier performance on emotional expression recognition, namely, average recognition rate, specificity (Spe), and Sensitivity (Sen). Sensitivity (Sen) and Specificity (Spe) refers to the efficiency of the classifier to detect a specific type of emotion as emotion and a particular class of emotions as other emotion, respectively.

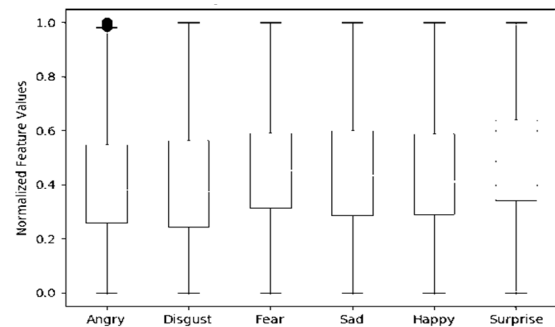


Fig. 4. Analysis of distance feature over six facial expression using Box plot.

III. RESULTS AND DISCUSSION

This section discusses the experimental results of the proposed methodology on facial emotional expression recognition. The proposed mathematical model-based algorithm placed the markers on a defined location in subjects face (Fig 3 (a)), and the optical flow algorithm efficiently tracked the movement of the marker. Figure 3 (b) shows the movement of the marker during happiness expression, and Figure 3 (c and d) shows the movement of markers in an optical flow algorithm. The distance feature

TABLE I. EMOTION CLASSIFICATION RATE USING KNN AND DT CLASSIFIERS (IN %)

Classifiers	Parameters		Fivefold cross-validation accuracy					Mean Classification Accuracy $\pm$ Standard deviation
	Distance Measure	K Value	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	
KNN	Manhattan	2	97.88	98.13	98.18	98.04	97.95	98.03$\pm$0.16
	Euclidean	6	97.44	97.34	97.87	97.22	97.44	97.45 $\pm$ 0.09
	Minkowski	3	96.92	97.76	97.13	97.27	97.01	97.24 $\pm$ 0.11
	Chebyshev	10	96.82	97.35	97.13	97.21	96.88	97.02 $\pm$ 0.20
Decision Tree			97.21	97.47	96.51	97.57	97.26	97.21 $\pm$ 0.18

TABLE II. INDIVIDUAL EMOTION CLASSIFICATION RATE USING KNN AND DT CLASSIFIERS (IN %)

Classifiers	Parameters		Individual Class Accuracy						Sensitivity	Specificity
	Distance Measure	K Value	Angry	Disgust	Fear	Sadness	Happiness	Surprise		
KNN	Manhattan	2	97.69	97.78	97.94	98.45	98.31	98.01	98.82	94.11
	Euclidean	6	96.95	97.15	97.27	97.84	97.95	97.51	98.46	92.35
	Minkowski	3	96.71	96.93	97.04	97.63	97.81	97.34	98.34	91.74
	Chebyshev	10	96.41	96.72	96.76	97.45	97.66	97.14	98.21	91.08
Decision Tree			96.75	96.79	97.22	97.74	97.64	97.14	98.32	91.65

extracted from 55 subjects is used for emotional expression classification using KNN, and DT classifiers. The experiment is conducted in a controlled environment with constant lighting and room temperature with different backgrounds. All the subjects are seated comfortably in a chair, which is 0.9 m from the camera system to capture the facial expressions. The proposed marker placement algorithm was tested under different head rotation angles and the results confirmed that, the placement of marker does not change up to 25° rotation in either clockwise/anticlockwise direction. However, all the subjects are instructed to restrict their head rotation during the emotional expressions until the end of the experiment. In this work, the subject's faces difficulties in perceiving and expressing disgust and sadness emotion through visual cues compared to other emotions. However, the box plot analysis of emotional expressions using normalized feature value in Fig 4 shows that all the emotions have a different mean value and could be distinguished from the acquired samples.

The mean emotional expression recognition rate of the KNN and DT classifier is given in Table I. The Manhattan distance measure in KNN gives a highest emotion recognition rate of 98.03% compared to DT classifier. In the case of the Manhattan distance measure, the lower value of k-neighbor gives the highest classification rate, and there is no significant deviation in accuracy has been found in a higher value of k-neighbors. But the accuracy of the emotion expression recognition rate increases as the number of neighbors increases in other three distance measures (Fig 5). Compared to KNN classifier, DT classifiers achieved a maximum emotion recognition rate of 97.21% and it does not perform-well in this study. This may be due to the size of the database. In this study, the number of features used for classification is more than 100000, and it may root the DT classifier in predicting the classes with higher accuracy.

However, the KNN classifier performs well for this data size and gives higher accuracy in lesser computation time because DT performs faster than KNN since the DT classifier does not have any external tuning parameters.

Table II shows the individual emotional expression recognition rate of two classifiers namely KNN, and DT. Among the four different distance measures, Manhattan distance measure gives the highest sensitivity of 98.82% and with a specificity of 94.11, compared to other distance measure. The Chebyshev distance measure gives the lowest mean sensitivity of 98.21%, and with a specificity of 91.08%. However, it's interesting to note that all the four-distance measure gives a higher sensitivity. Here, negative emotions such as angry, disgust, and fear emotions gives lower accuracy compared to positive emotions (happiness, and surprise).

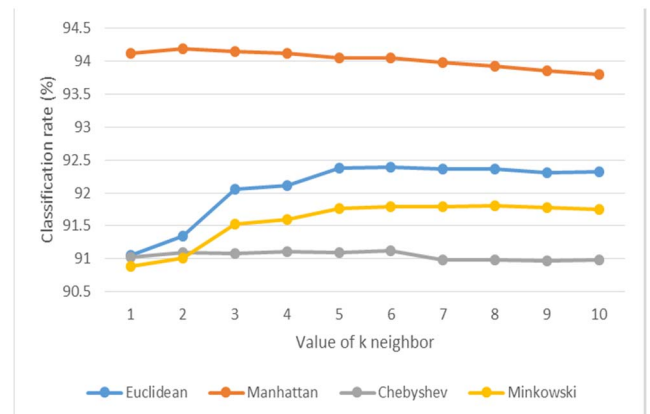


Fig. 5. Mean accuracy of KNN classifier over four distance measures

TABLE III. COMPARISON OF EMOTIONAL EXPRESSION RECOGNITION RATE WITH EARLIER WORK

Ref	Classifier	Network parameter		Angry	Disgust	Fear	Sadness	Happiness	Surprise	Sensitivity	Specificity
[14]	PNN	$\sigma=0.01$		92	90	89	88	96	94	98	90
	ELM	MLP random layer	tanh	90	88	85	84	92	91	97	86
Present work	KNN	Manhattan	K=2	97.69	97.78	97.94	98.45	98.31	98.01	98.82	94.11

In KNN, the accuracy of emotions also depends on the value of K. Here, the lowest value of K gives a higher emotion recognition rate. The main reason behind is lower accuracy is due to perceptual ability and expressiveness of the subject during the experiment. Most of the subjects could not accurately and naturally express these emotions as like other emotions (happiness, and surprise). The performance of the proposed facial emotion expression recognition system is comparable to the earlier works in the literature. In our previous work, we used PNN and ELM classifiers for classifying facial emotional expression using distance features and achieved a maximum mean classification rate of 94% and 92%, respectively [14]. But in the present study using KNN, and DT classifiers, the maximum mean accuracy is >97%, and comparable with state-of-the-art methods in the literature. However, the work aims to analyze different types of features besides distance features, and diverse machine learning, and deep learning algorithms (CNN and LSTM) for emotional expression classification in the future.

The distance features based facial expression recognition achieved a maximum sensitivity, and specificity of 98.82% and 94.11% using KNN classifier with Manhattan distance measure. However, the accuracy of the facial expression recognition using the proposed methodology could be improved in the future through the following works: (a) the proposed facial emotion recognition algorithm is only tested with our facial expression database. In future, the proposed methodology on facial emotion recognition could consider the international standard facial expression databases for improving its efficiency and robustness in facial emotion recognition. (b) different types of facial geometric features based on FAUs movements over six emotions could be investigated in the future for improving the accuracy of facial emotion recognition besides a simple distance measure used in the present work. (c) different types of nonlinear classifiers such as ensemble classifiers, Support Vector Machine (SVM), Neuro-fuzzy classifiers, and Random forest classifiers could be investigated in the future for improving the accuracy of emotion classification using facial geometric features.

IV. CONCLUSION

In this work, ten facial action units are used to classify six basic emotional expressions. The distance between each FAUs to the center of the face marker is used as features to classify the emotional expression using two classifiers namely, Decision Tree (DT), and K Nearest Neighbor (KNN). The FAUs are placed on the subject's face at a defined location using the proposed mathematical model, and movement of markers during the emotional expressions is computed through Euclidean distance measure. The distance features extracted from six emotions are tested for its significance using one-way analysis of variance

(ANOVA) with a significant level of $p < 0.01$. The experimental results confirm that the distance feature is highly significant in classifying facial expressions. In a KNN classifier, the value of closes neighbor (k) is varied

between 1 and 15, and the performance of the classifier is assessed through a grid search method, and default parameters are used in a DT classifier for emotional expression classification. The proposed methodology achieved a maximum mean emotional expression rate of 98.03% and 97.21% using KNN, and DT classifiers, respectively, using a fivefold cross-validation method. The accuracy of emotional expression achieved on this work is high compare to the state of the artworks in the literature, and the computational complexity analysis (execution time and memory requirement) is less. Besides, the proposed system also facilitate the user to develop a real-time emotional expression detection system for different applications such as biometric, bipolar disorder assessment, smart home design, etc.

ACKNOWLEDGMENT

The authors would like to acknowledge and thank the Kuwait Foundation for the Advancement of Sciences (KFAS) for the financial support through an exploratory research grant. (No: PR18-18EE-01).

REFERENCES

- [1] S Macros, E Gonzalez, C Martin, L Flores, J Gomez, G Bermejo, E Zalama, "Virtual Avator for Emotion Recognition in Patients with Schizophrenia: A Pilot Study", Fron in Human Neuroscience, 2016. DOI: 10.3389/fnhum.2016.00421.
- [2] F Smith, S Rossit, "Identifying and detecting facial expressions of emotion in peripheral vision", PlosOne, 13(5): e0197160. <https://doi.org/10.1371/journal.pone.0197160>
- [3] Redferm, "Emotion Tracking for Remote Conferencing Applications using Neural Networks", 21st National Conference on Artificial Intelligence and Cognitive Science, pp, 1-6, 2010.
- [4] T Teng, X Yang, "Facial expressions recognition based on convolutional neural networks for mobile virtual reality", 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry, Vol 1, 475-478, 2017.
- [5] C Pelachaud, "Modelling multimodal expression of emotion in a virtual agent", Phil. Trans. R. Soc, 364, pp, 3359-3348, 2009.
- [6] A Mahmood, S Hussain, K Iqbal, W Elkilani, "Recognition of Facial Expressions under Varying Conditions using Dual-Feature Fusion", Math. Prob. In Engg, 2019, 1-12. DOI: 10.1155/2019/9185481
- [7] R E Hamer, E Rusadze, I Lusi, T Ahmad, S Escalera, G Abarjafari, "Review on Emotion Recognition Databases", Intech, 2017. DOI: 10.5772/intechopen.72748.
- [8] L Feldman, R Adolphs, S Marsella, Aleix M, S Pollak, "Emotional Expressions Reconsidered: Challenges to Inferring Emotions from Human Facial Movements", Physc. Sci. in the Pub Int, 20(1), 1-68, 2019.
- [9] Z Yu, G Lu, Q Liu, J Deng, "Spatio-temporal convolutional features with nested LSTM for facial expression recognition", Neurocomputing 317, 50-57, 2018.

- [10] V Maruthapillai, M Murugappan, "Optimal Geometric Set for Automated Marker Placement to Virtualized Real-Time Facial Emotions", PlosOne, 11(2), e0149003, 2016. DOI: 10.1371/journal.pone.0149003.
- [11] A Sherly, D Dharma, "Novel directional patterns and Generalized Supervised Dimension Reduction System for facial emotion recognition", Mult. Tools. and Appl, 2017. DOI: 10.1007/s11042-017-5141-8.
- [12] Open Computer Vision Library, www.opencv.org
- [13] V. B. Surya Prasath, H A A Alfeilate , A B. A. Hassanate , O Lasassmehe, A S. Tarawnehf, M B Alhasanatz, H S.E Salmane, Effects of Distance Measure Choice on KNN Classifier Performance - A Review, arXiv:1708.04321v3, 2019.
- [14] F Nonis, N Dagnes, F Marcolin, E Vezzetti, 3D Approaches and Challenges in Facial Expression Recognition Algorithms – A Literature Review, Applied Sciences, 9 (3904), 2019. DOI: 10.3390/app9183904.
- [15] M Murugappan, R Nagarajan, S Yaacob, "Combining spatial filtering and wavelet transform for classifying human emotions using EEG signals", Journal of Med. and Bio Engg, 31(1), 45-51, 2011.
- [16] R Yuvaraj, M Murugappan, M Norlinah, K Sundarj, M Khairiyah, "Review of emotion recognition in stroke patients", Dementia and Geriatric Cognitive Disorders, 36(3-4), 179-196, 2013.
- [17] P Quelhas, J Stoyanovam Marker versus Markerless Augmented Reality. Which Has More Impact on Users?, Int. Jou. of Hum. Comp. Inter, 34(9), 819-833, 2017.