Recognition of self-initiated emotional expressions using facial electromyography

Vladislav Aksiotis
Center for Bioelectric interfaces, Institute for Cognitive Neuroscience, HSE University
Moscow, Russia
v.aksiotis@gmail.com

ABSTRACT

A considerable body of research exists concerning the highly precise classification of facial expressions through the utilization of facial electromyography (fEMG). These studies typically involve instructing participants to display specific expressions. However, this structured approach does not mirror the spontaneous nature of emotional expressions in real-world communication. Furthermore, the computational load of the common methods used in these studies often renders them impractical for online application.

In contrast, our study investigates the feasibility of classifying a broad array of spontaneous, self-initiated emotional facial expressions such as happiness, anger, and more. We encouraged participants to freely choose the way to demonstrate emotions. The emotions represented included both basic and complex ones, encompassing twelve distinct categories in total.

We employed eight optimally positioned myographical electrodes on the face for the capture of expression data. Each participant was tasked with displaying each emotional pattern thirty times, with each instance lasting three seconds to create a sufficiently diverse sample set.

To classify these expressions, we extracted three principal features: the mean amplitude of the signal envelope, as well as the first and second peaks of frequencies following the Fourier transformation. The primary classification method used was the random forest technique.

Results from ten participants demonstrated an average recognition accuracy of 70% for the distinct facial patterns. This accuracy was maintained even when the model was trained on one subset of participants and tested on another. We found the electrodes placed on the middle of the forehead and the cheeks to provide the most informative data. Topographies of significance for each expression are also presented.

Our findings suggest the potential for the proposed EMG technique to facilitate communication of emotions in scenarios where visual cues are absent, such as in virtual reality technologies. Moreover, the simplicity of the data processing makes real-time, online classification feasible, providing a foundation for further research into the automation of emotion recognition using sEMG.

KEYWORDS

Facial expressions recognition, facial expression, emotions, classification, electromyography

1 INTRODUCTION

The development of an accurate emotion classifier, leveraging various neuroscience methodologies, stands as a fundamental objective within the field of emotional artificial intelligence. Research has been largely focused on constructing facial expression classifiers utilizing data captured through facial electromyography (EMG). Inspired by these efforts, we set out to advance the field by creating our own iteration of such a procedure, capable of capturing and categorizing not just basic emotions but also a selection of more nuanced, non-basic ones.

Our approach was designed with the explicit aim of showcasing the potential for successful recognition of a wide spectrum of emotional expressions. In addition to its expansive scope, it is also noteworthy for its simplicity in computational requirements and its minimalistic utilization of hardware. Specifically, it hinges on the extraction of straightforward yet insightful derived features from a limited number of optimally placed electrodes. This balance between complexity and practicality is intended to bolster the accessibility and applicability of this emerging technology, thus extending its reach and impact in the field of emotional AI.

2 BACKGROUND

The field of emotion recognition and classification, particularly those related to facial expressions, has seen an extensive array of studies leveraging EMG [1, 8, 9, 10]. Among the seminal findings from the 1980s was the fundamental validation of EMG's efficacy for emotion detection [13], marking a significant milestone in the field.

The utility of facial electromyography (fEMG) for detecting and identifying emotions such as anger, sadness, fatigue, and pleasure has been robustly established in subsequent research [5]. The technology's application extends beyond the realm of mere detection, and it has been used effectively for the classification of expression patterns, demonstrating an impressive accuracy range from 0.9 to 1 [10, 8].

Exploring the physiological underpinnings, certain studies have noted an increase in the corrugator supercilii muscles (surrounding the eyes) activity in response to expressions of anger and sadness, while the opposite effect was observed for expressions of happiness [6]. It's speculated that this could be attributed to the less cognitively demanding task of identifying happy faces compared to their more emotionally charged counterparts. Moreover, specific patterns for the expression of disgust, including the activation of the M. corrugator and M. orbicularis oculi muscles, have been found to distinguish it distinctly from expressions of attraction and joy [7].

With the advent of technological advancements, virtual reality has emerged as a promising platform for optimizing human-computer interaction [8]. Facial expressions, assumed to mirror the emotional states, are inferred by observing changes in muscular electrical activity. In a significant contribution, Emteq labs (2020) [2, 3] have developed a tool that allows emotion analysis within a virtual reality headset. This technology employs multi-modal biometric sensors to track variables like heart rate, fEMG, skin conductance, eye and body movements, offering a comprehensive measure of emotional states. Despite the potential of this technology, its practical applications may be limited, particularly when the objective is interpersonal communication.

A notable gap in previous research pertains to the focus on precisely defined patterns or emotional states. Each individual, however, possesses a unique manner of emotional expression. In this study, we propose an innovative approach, concentrating on self-induced expressions. We explore the feasibility of generalizing and predicting emotional patterns using a maximum of eight myographical electrodes, contributing a nuanced understanding to the field.

3 METHODS

The present study explores the recognition of both "basic" and less common emotional facial expressions utilizing machine learning methods for classification. The selection of emotions incorporated into our study leans on the frameworks provided by Paul Ekman [11] and Robert Plutchik's theory of emotions [12]. The total number of emotions under study amounted to 11, with an additional baseline state (neutral) which the classifier should not identify as an emotional pattern.

Overall, basic emotions were anger, disgust, fear, sadness, happiness, surprise. The rest were anxiety, contempt, delight, perplexity, and pride.

3.1 EMG Setting

Our method involved the use of fEMG for data collection. The data were captured using eight electrodes, coupled with electroconductive gel. Each side of the participant's face had three electrodes placed over the skin overlying the occipito-frontal, zygomaticus major, and zygomaticus minor muscles. Two further electrodes were located above the procerus muscle and mentalis.

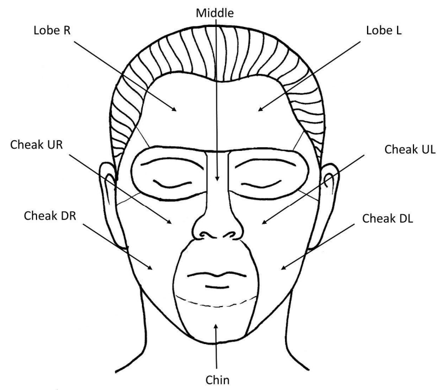

Figure 1. Scheme of electrodes positions. Adapted from [14].

For reference and grounding purposes, an additional two electrodes were placed on the left side of the participant's neck and left collarbone, respectively.

3.2 Experimental procedure

The participants were briefed about the study's objectives and given clear instructions regarding their roles. Upon receiving informed consent, the electrodes were affixed to their faces and necks as previously outlined. An essential aspect of our study was the self-initiation of emotional expressions rather than inducing them externally. Participants were requested to express a series of 11 emotions and neutral state as naturally as possible. The choice and manner of emotional expression were left to the participant's discretion.

Each of the 12 emotional expressions was recorded over 30 trials, with each trial lasting for three seconds and a two-second interval between samples. The total duration of the experiment ranged between 35-40 minutes for each participant. The recording's sampling frequency was set at 1000 Hz. Our study's sample was composed of 10 volunteers (3 males), aged between 18 to 23 years (M: 20.3, SD: 1.5).

3.3 **Data analysis**

The analysis of our data involved a feature extraction approach applied to the signals from each trial. Before feature extraction, every channel was processed using a bandpass Butterworth filter (ranging from 10 to 350 Hz) to eliminate non-target frequencies. Following this, the Fourier transform was used to extract two peak frequencies from the signals. Then, we calculated the signal's envelope by applying the Savitzky–Golay filter to the rectified (absolute) values. The resulting average and maximum amplitude of the envelope were included as features in our data set. Figure 2 visually represents the generalized scheme for feature extraction. As a result, average and maximum amplitude and two frequencies' peaks in every electrode were used to train models since these features showed significant impact on the accuracies. Overall, dataset for every person had size 360 (samples) by 32 (8 electrodes * 4 features).

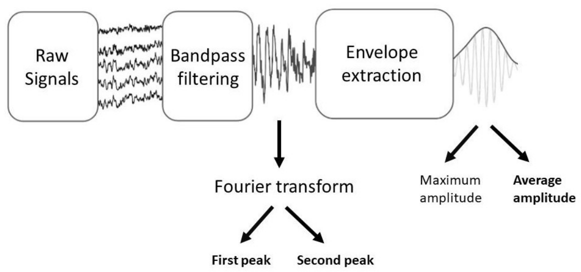

Figure 2. Scheme of the raw signal analysis and feature extraction. Bold text indicates the most important features.

From the diverse range of available classification methods, we selected decision trees, random forests, and gradient boosting as potential algorithms for classification. The random forest algorithm yielded the best results and was thus used as the final model. Models were trained and tested using the scikit-learn package [4].

We also explored two designs for model training and testing. The first design involved blending the general dataset and then randomly dividing it into training and testing sets at a 3:1 ratio. This approach is only feasible when data can be collected from participants and incorporated into a retrained model. The second design segmented data by participants, meaning the models were trained on a larger group and then tested on a separate participant. We propose that this approach could be suitable for real-time systems as it eliminates the need to train the model from scratch.

4 **RESULTS**

Our first set of results pertains to the classification outcome derived from the initial training approach for the models, where the data was mixed. The overall accuracy of recognition, combining all 8 electrodes, reached 0.85 (SD: 0.05), in contrast to a 0.08 probability for random guessing. The average F-measure across all classes mirrored this accuracy. The corresponding metrics, illustrated in Figure 3a, show these results graphically.

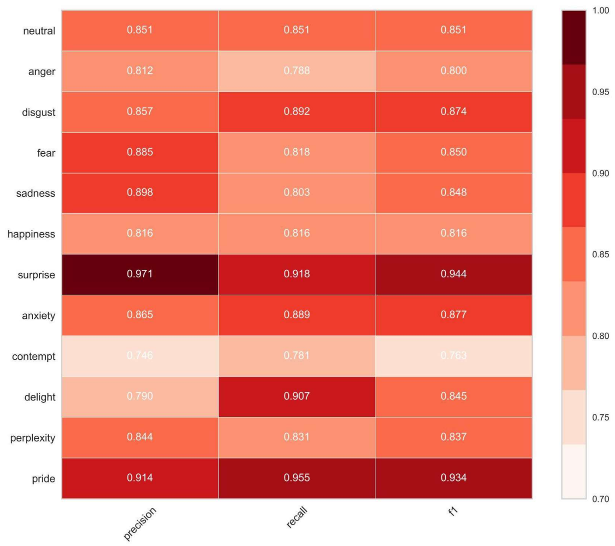

Figure 3a: The resulting metrics for classification of the 12 expressions.

In terms of individual emotions, the model achieved the highest recognition results for expressions of happiness and neutral states. When considering all 12 expressions, the emotions of anxiety and perplexity had the lowest accuracy of recognition. We then created a separate model that focused solely on basic emotions, which delivered an overall accuracy of 0.89 (SD: 0.07), with a 0.14 probability of random guessing. These metrics are depicted in Figure 3b.

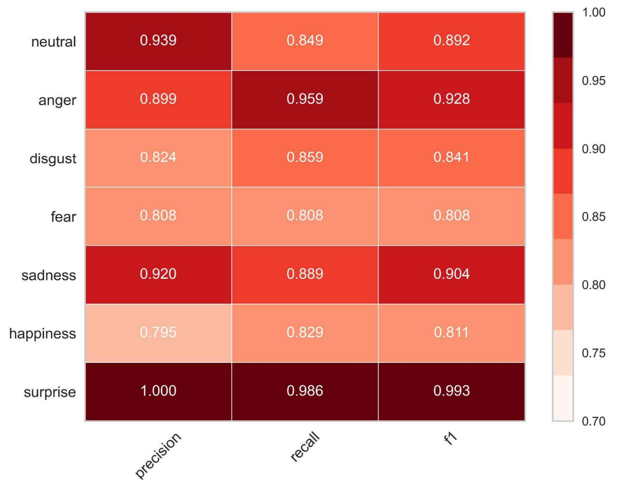

Figure 3b: Metrics for 6 expressions for basic emotions.

Our analysis then transitioned to the second training approach, which involved splitting the training and testing samples based on individuals. After splitting the data were standardized in order to eliminate individual differences and make it possible to predict a participant's expressions by features of other people.. As shown in Tables 1 and 2, this change in approach resulted in a notable decline in classification accuracy. The average accuracy for all expressions fell to 0.29 (SD: 0.06), and for basic emotions, it stood at 0.58 (SD: 0.13).

Subject ID	1	2	3	4	5	6	7	8	9	10	Mean	SD
Accuracy	0,31	0,38	0,26	0,34	0,14	0,29	0,31	0,28	0,31	0,25	0,29	0,06
Recall	0,31	0,39	0,25	0,33	0,13	0,3	0,23	0,27	0,28	0,23	0,27	0,07
F-score	0,34	0,42	0,27	0,34	0,16	0,31	0,27	0,27	0,32	0,22	0,29	0,07

Table 1: Classification accuracy on individual participants as test samples for the full range of expressions.

Subject ID	1	2	3	4	5	6	7	8	9	10	Mean	SD
Accuracy	0,69	0,69	0,76	0,69	0,35	0,56	0,61	0,48	0,53	0,43	0,58	0,13
Recall	0,64	0,66	0,71	0,66	0,31	0,55	0,56	0,46	0,53	0,39	0,55	0,12
F-score	0,63	0,7	0,72	0,66	0,34	0,55	0,57	0,49	0,6	0,38	0,56	0,12

Table 2: Classification accuracy on individual participants as test samples (only over the basic emotions expressions). Subject ID - which data were used for testing.

We also conducted an analysis of pairs of expressions using binary classification. This was to gain deeper insights into which patterns are similar in terms of muscle contraction. The results are depicted in Figure 4.

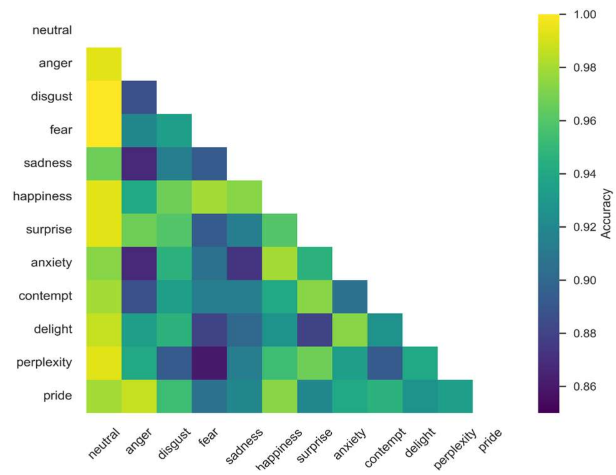

Figure 4: Average accuracy of binary classification between pairs of expressions.

The most frequent misclassifications occurred between the following pairs: Perplexity and Fear, Surprise and Fear, Anger and Disgust, and Surprise and Delight.

In addition to these, we also analyzed pairs of electrodes separately, with models trained on mixed data. Based on these results, the highest average accuracy, at 0.55, was achieved with two electrodes located on the left cheek and above the nose. Combinations of electrodes placed on the chin muscles, zygomatic muscles (lower part of the cheek), and procerus (above the nose) produced the highest accuracy rates. Figure 5 provides a graphical representation of these accuracies.

Figure 5: Average accuracy of binary classification between pairs of electrodes.

Topographies of emotional expressions can be extremely useful when deciding on which position to put electrodes or other devices for measuring physiological processes to monitor target expression. To create the topography, we used logistic regression between the neutral state and the target expression on a generalized dataset that was standardized within the participants. Next, we got the weights of the models and used them to display on the face diagram. The topographies are shown in Figure 6.

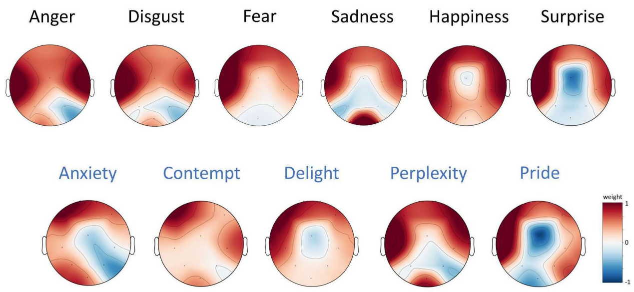

Figure 6. Averaged expressions topographies. Color - the weights (significance) of the space for separating neutral and target expression. Positive values correspond to a larger amplitude and frequency, negative values correspond to smaller ones.

Topographies can be an important factor for the separation of expressions. For example, comparing Figure 4 and topography, it can be seen that anger, which is frequently misclassified between anxiety and sadness, has common features of topography with these expressions.

5 DISCUSSION

Our study yielded encouraging outcomes that contribute valuable insights into the variety of self-induced emotional expressions. One critical finding that emerges from the data is the consistent understanding of basic emotions across different participants, as evident from the accuracy levels of models trained on data from various participants.

In terms of specific emotions, the relative difficulty in differentiating between anxiety and perplexity could be attributed to individual variances in how these emotions are mentally represented and their contextual similarity in day-to-day situations. Moreover, the emotional expressions of fear, surprise, and perplexity seem to share a similar pattern that can be seen by topographies. The participants tended to open their mouths and widen their eyes during the display of these emotions. To accurately distinguish between these expressions, additional information pertaining to physiological responses or situational context may be required.

When evaluating the classification of emotional expressions across different individuals, we observed that the resulting models exhibited relatively low generalizability. The high accuracy rates attained from models trained on mixed data likely stem from the models' effectiveness in accurately identifying patterns within each individual, subsequently boosting overall accuracy. However, when we split the data by participant, the basic emotional expressions were classified with significant accuracy. These results support the idea that within a specific culture, there are common mental representations of familiar emotional expressions. Consequently, this understanding could contribute to the development of an emotional communication system that does not necessitate additional computational resources for model training.

However, there are limitations to our study that warrant attention. A significant constraint is the cultural homogeneity of the participants, who are solely of European descent. In addition, the sample comprised mostly females, which further restricts the scope of our results and amplifies the homogeneity of our sample.

6 CONCLUSION

The findings of our study indicate that it is feasible to develop a computationally efficient classification procedure with high recognition rates for both basic and atypical facial expressions. Despite a slight reduction in accuracy, pattern recognition is achievable using a pair of electrodes. Even though our models made errors in classifying certain pairs of expressions that bear similarity to each other, their classification can still be deemed reasonably accurate.

A significant hurdle we encountered in terms of the model's generalizability was a decrease in overall accuracy. It's crucial to highlight this limitation as our proposed model can only classify expressions from new individuals within a restricted set of expressions. However, if the training set can be augmented with new data, more favorable results could be achieved.

The resulting topographies indicate a possible reason for the decrease in the accuracy of the model within some expressions, since many emotional expressions have common myographic patterns, especially on the cheeks.

We envision that our approach could be practically implemented to swiftly recognize emotional expressions through the recording of facial electromyography. Such an application could prove beneficial in domains such as virtual reality technologies and emotional AI.

REFERENCES

[1] Arturo Nakasone, Helmut Prendinger, Mitsuru Ishizuka. 2005. Emotion recognition from electromyography and skin conductance. *In Proc. of the 5th international workshop on biosignal interpretation*, 219-222.

[2] Emteq labs. Retrieved from https://www.emteqlabs.com/.

[3] Emteq labs. Measuring what Matters in Immersive Environments. Retrieved from https://www.emteqlabs.com/wp-content/uploads/2020/11/Measuring_what_Matters_in_Immersive_Environments.pdf

[4] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Andreas Müller, Joel Nothman, Gilles Louppe, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. 20111. Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*, 12, 2825-2830.

[5] Huang Cheng-Ning, Chen Chun-Han, Chung Hung-Yuan. 2004. The review of applications and measurements in facial electromyography. *Journal of Medical and Biological Engineering*, 25(1), 15-20.

[6] Janina Künecke, Andrea Hildebrandt, Guillermo Recio,Werner Sommer, Oliver Wilhelm. 2014. Facial EMG responses to emotional expressions are related to emotion perception ability. *PloS one*, 9(1), e84053, doi: https://doi.org/10.1371/journal.pone.0084053.

[7] Karsten Wolf, Reinhard Mass, Thomas Ingenbleek, Falk Kiefer, Dieter Naber, Klaus Wiedemann. 2005. The facial pattern of disgust, appetence, excited joy and relaxed joy: An improved facial EMG study. *Scandinavian journal of psychology*, 46(5), 403-409, doi: https://doi.org/10.1111/j.1467-9450.2005.00471.x.

[8] Lorcan Reidy, Dennis Chan, Charles Nduka, Hatice Gunes. 2020. Facial Electromyography-based Adaptive Virtual Reality Gaming for Cognitive Training. *In Proceedings of the 2020 International Conference on Multimodal Interaction*, 174-183, doi: https://doi.org/10.1145/3382507.3418845.

[9] Louisa Kulke, Dennis Feyerabend, Annekathrin Schacht. 2020. A comparison of the Affectiva iMotions Facial Expression Analysis Software with EMG for identifying facial expressions of emotion. *Frontiers in Psychology*, 11, 329, doi: https://doi.org/10.3389/fpsyg.2020.00329.

[10] Mahyar Hamedi, Sh Hussain Salleh, Chee Ming Ting, Mehdi Astaraki, Alias Mohd Noor. 2016. Robust facial expression recognition for MuCI: a comprehensive neuromuscular signal analysis. *IEEE Transactions on Affective Computing*, 9(1), 102-115, , doi: 10.1109/TAFFC.2016.2569098.

[11] Paul Ekman. 1999. Basic emotions. *Handbook of cognition and emotion*, 98(45-60),16, doi: https://doi.org/10.1002/0470013494.ch3.

[12] Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. *Theories of emotion*. Academic press, 3-33.

[13] Ulf Dimberg. 1988. Facial electromyography and the experience of emotion. *Journal of Psychophysiology,* 2(4), 277–282.

[14] I. Gerych, V. Savchyn, I. Stojanowski, A. Barwinska, O. Chemerys. 2013. New classification of head burn injuries. Actual problems of modern medicine: Bulletin of the Ukrainian Medical Stomatological Academy. 13. 206-210.