2024 18th International Conference on Automatic Face and Gesture Recognition (FG)

The Seven Faces of Stress: Understanding Facial Activity Patterns during Cognitive Stress

Carla Viegas[1,3], Roy Maxion[2], Alexander Hauptmann[1], and João Magalhães[3]
[1] Language Technologies Institute, Carnegie Mellon University, Pittsburgh, USA
[2] Computer Science & Machine Learning, Carnegie Mellon University, Pittsburgh, USA
[3] NOVA LINCS, Universidade NOVA de Lisboa, Lisbon, Portugal

***Abstract*— Stress has been recognized as one of the main contributors to mental health problems, as well as cardiovascular diseases. To reduce the risk of severe diseases, early detection of stress is needed. One of the recent methods studied to detect stress is through facial expression analysis from videos. Although computer vision techniques combined with deep learning have been shown to detect stressful faces, there is a lack of work attempting to define how stressful faces look. One of the main challenges is that the expression of stress is person-dependent and one individual can show stress in various ways. In this work, we present a semi-automatic method that allows to distill from a large quantity of data facial activity patterns that are recognized to show stress. We are the first to combine quantitative and qualitative methods on data from 115 subjects to identify and propose seven facial activity patterns during stress. We support this proposal by analyzing the relationship of the different stress facial expressions with the basic emotions and show how individual components of anger, fear, surprise, and sadness co-occur during our defined stress facial activity patterns.**

Fig. 1: Examples from our seven different stress facial activity patterns defined through semi-automatic methods. We named the clusters as follows (from top left to bottom right): "*pressed lips*", "*biting lip*", "*stoic face*", "*open mouth*", "*lifted eyebrows*", "*frowning*", and "*dimpler*". All samples used to define the seven facial activity patterns of stress were identified by user study participants as showing stressed individuals.

I. INTRODUCTION

Stress is a natural reaction to danger, causing a spike of adrenaline in the bloodstream which sharpens our senses and decreases reaction time. Nature's purpose of stress is to increase the probability of surviving life-threatening situations. While stress used to save our lives in punctual encounters with dangerous animals, stress has become for a majority of society an ever-present companion due to financial worries, never-ending to-do lists, and continuous availability. Continuous exposure to stress is extremely harmful to the human body and has been shown to contribute to cardiovascular diseases which is the number one cause of death in Western countries. Especially in the workplace, stress can be prevented and treated, allowing workers who receive treatment to be more productive [7].

But how can stress be measured in a non-invasive manner? Clinical methods are mostly invasive and involve blood analysis and the use of electrocardiography (ECG) to compute accurately the heart rate-variability (HRV). With the development of smart wearables, sensors have been developed to evaluate electrodermal activity (EDA) and HRV. Several authors [17], [33], [34] studied the capabilities of using HRV measurements of the Apple Watch to distinguish between mental stress and relaxation, however, they come to different conclusions. Another non-invasive method to detect stress is the analysis of facial expressions. Viegas et al. [35] has shown for the first time that facial action units (AUs) defined by Ekman et al. [11], can distinguish between stress and non-stress activities. Other authors [15], [31], [29] have confirmed the effectiveness of AUs for stress detection. Facial analysis to detect stress is especially useful in working environments where webcams are already in place or easily installed. In stressful work settings, these applications could not only lead to more timely and reduced-cost interventions [19] but also to more productive environments where workers could better manage their workload [1].

Although the interest in using AUs for stress detection has grown in the last years, a database with prototypical stressed faces similar to what exists for basic emotions, is still missing. Facial expressions during stress seem not to be universal, but rather person-specific [35], [23], [8], [5]. Studies using AUs for stress detection, such as [35], [15], [16], [14], [5] lack a deeper analysis that includes a user study recognizing stress from the facial expressions as well as an analysis showing the relationship with emotion-related AUs for the different facial activity patterns that occur during stress.

Several challenges exist in defining facial activity patterns

This work has been partially funded by the FCT project NOVA LINCS Ref. UIDP/04516/2020.

979-8-3503-9494-8/24/$31.00 ©2024 IEEE

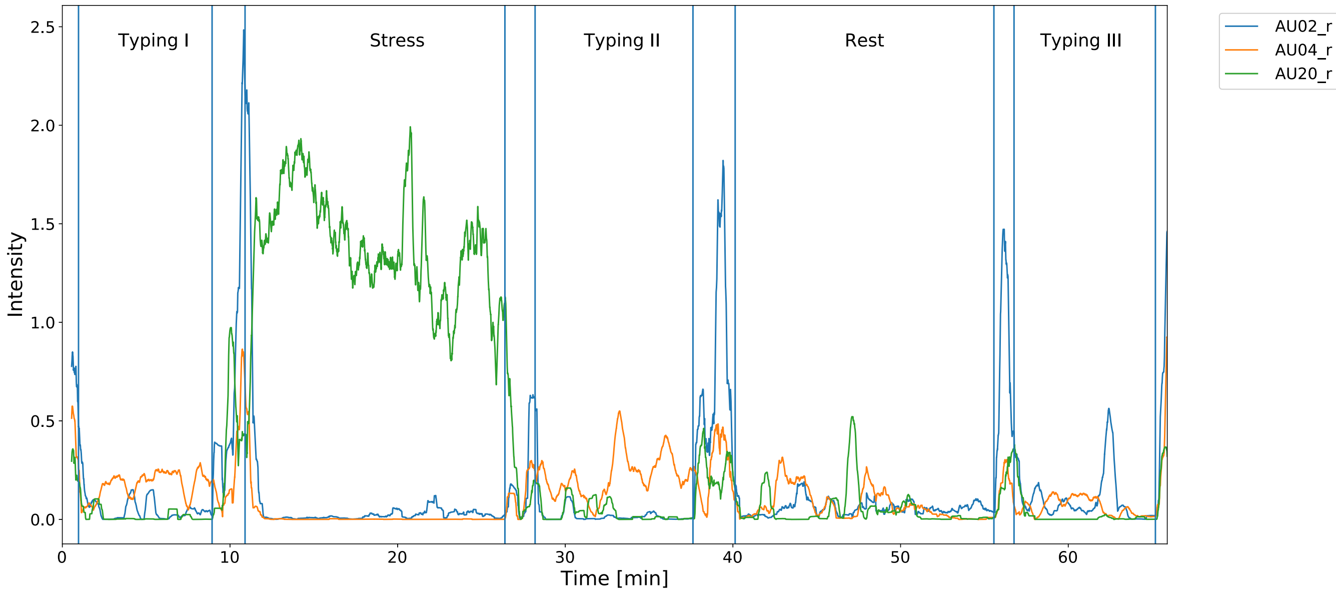

Fig. 2: Plot shows the activity of AU2 (outer brow raiser), AU4 (brow lowered), and AU20 (lip stretcher) of one subject during the entire experiment. The start and end of each experiment phase are marked by vertical lines. In between the main phases, the subjects were given short breaks, during which, we can observe an increased intensity of the outer brow raiser. The plot shows an increased intensity of the lip stretcher only during stress.

of stress: 1) facial expressions of stress are person-specific, 2) one individual could show stress through different facial expressions, 3) the human face allows for an immense number of facial expressions, 4) we do not know which characteristics of facial expression are relevant during stress, nor 5) do we know how many facial activity patterns we are looking for.

In this work, we present an analysis methodology that figuratively speaking, allows us to find several needles in a haystack. We analyzed a total of 575 minutes of video recordings from 115 subjects, combining clustering with human annotation to define facial activity patterns that show stress. Our methodology does not require trained Facial Action Coding System (FACS) specialists or psychologists for data annotation. Instead, we employed automatic AU detectors, computed clusters from a large amount of data points, and used cluster centroids for human annotation. With our proposed method, we identify seven facial activity patterns during stress. As facial expressions of stress are often related to disgust and anger [13], we show the relationship between active AUs during stress and the basic emotions. With this work, we aim to provide a novel resource to recognize stress and simulate stress-related facial behavior to make human-machine interactions more natural. For this purpose, characteristic facial activity patterns during stress are made publicly available[1].

II. Related Work

The study of stress began during a period in which experimental psychology was still in the shadows of behaviorism (1900-1980), which assumed that the mind was just a black box between stimulus and response, not worthy to be studied [22]. Richard Lazarus played a major role in shifting the field of psychology and 1984 proposed the "Transactional Model of Stress and Coping" which emphasizes the person–environment transaction and suggests that a stress response is highly influenced by individual appraisal processes [22], [21].

In parallel Ekman et al. [10] had already defined the Facial Action Coding System (FACS) which is a taxonomy for facial muscle movement and studied facial action units (AUs) during stress-induced situations. Based on their theory of basic emotions, they found the emotion of disgust to co-occur during stress. Since then the relationship between facial expressions and stress has been studied from different angles. In [26], neurobiologists studied how gender influences the behavioral stress response, finding that men show greater stress-induced corrugator reactivity (frowning in the absence of any stimulus) than women. In [36] evolutionary psychologists found that displacement behaviors which are known to be associated with stress influence the likability of a person, indicating a benefit and potential adaptive function of displaying stress through facial expressions.

Detecting stress from facial expressions has also come into the focus of computer vision. The initial studies focused on detecting emotions such as disgust and anger to assume the presence of stress [13]. Viegas et al. [35] were the first to use FACS and train different machine learning algorithms to detect stress solely from facial expressions. They showed that subject-dependent models were able to detect stress with an accuracy of up to 91% while subject-independent

[1] https://github.com/clviegas/SevenFacesOfStress

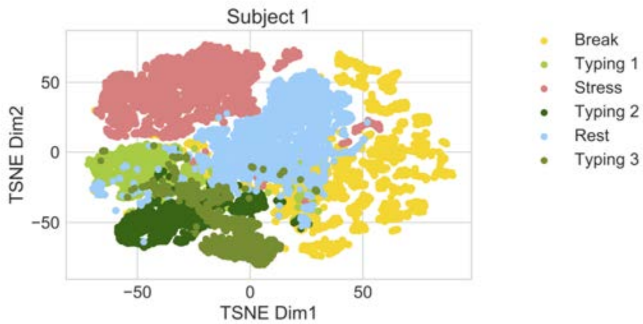

Fig. 3: Visualization of the AU intensities of one subject during the entire experiment, using t-SNE. Each feature point corresponds to one frame. The different colors represent the different experiment phases.

models reached an accuracy of 74% using Random Forest. Since then, different researchers created their own datasets to train machine learning models for recognizing stress from AUs [15], [31], [29]. Unfortunately, the datasets are not public.

Creating high-quality stress datasets requires a carefully designed experimental protocol. [15] collected data from 24 participants who underwent different stress-inducing activities such as social exposure, emotional recall, and mental workload. [29] and [5] collected data from 40 and 62 participants respectively during the Trier Social Stress Test also collecting neuroendocrinological stress levels, such as saliva cortisol. [31] on the other hand created a dataset of 34 subjects participating in online video meetings with self-reported stress levels. To the best of the authors knowledge, the video dataset used in this work is unique in the size of over 110 participants and in the careful experimental design including HRV and BP in addition to self-reported stress levels.

The features used in previous work studying AUs ([31], [29], [5], [35]) have been mainly obtained with the OpenFace toolbox [3]. [15] on the other hand, trained their own AU detector with deep learning methods. While [31] and [15] focus in their work on obtaining high classification accuracies for stress, [29] studies the differences in the performance of ML models depending on the type of stress labels used, e.g. self-reported, live-observed, video-annotated, and neuroendocrinological. Besides obtaining an accuracy of 81.1% with an SVM, [15] performed the nonparametric Wilcoxon signed-rank test on the AUs, concluding that all AU intensities are higher during stress. [5] on the other hand, focused solely on statistical tests showing the quartiles of each AU during different experimental phases. Also, AU occurrence was analyzed, concluding that AU 5 (upper lid raiser), 7 (lid tightener), and 10 (upper lip raiser) occur more often during stress. So far, our community lacks research in which facial activity patterns are analyzed and defined, moving beyond ML classification and statistical analysis.

III. Data

In this work, we utilize the video recordings of the dataset created by Lau [18] which is composed of data from 115 subjects (48 male and 67 female). The original goal of the dataset was to determine if detecting stress through keystroke dynamics is possible. Nevertheless, frontal video recordings of the entire experiment, as well as, qualitative and quantitative evidence of actual stress state during the stressor task, make the dataset a unique resource to study facial expressions during stress.

A. Recording Setup

The primary experiment detailed in [18] involved capturing four distinct video streams. Three of these streams exclusively captured the subjects' keyboard typing activities from different angles: left, right, and above. The fourth stream featured a frontal view capturing the subject's face during the experiment, which is the specific video data analyzed in this work. The recording equipment was a Microsoft Life Studio Pro webcam with 1080p resolution, capturing at 30 frames per second. The frontal camera was positioned beneath the monitor utilized by participants throughout the experiment. Each video has an approximate duration of one hour, with slight variations depending on the typing speed of the subject.

B. Experimental Protocol

The facial video data formed a component of a broader experiment with the primary objective of gathering typing data from subjects under both neutral and stressed conditions. The overarching aim was to determine the feasibility of discerning between subjects' typing in neutral and stressed states. Alongside typing data, additional ancillary information such as BP and ECG readings was collected to assess participants' stress levels induced by the stressor.

All participants in the experiment followed the same protocol. Initially, each subject underwent a 30-minute rest period to establish a neutral baseline. Next, the subject provided a neutral typing sample ("*Typing 1*"), capturing the initial neutral face video. Subsequently, the subject engaged in a 15-minute stressor task, involving a multitasking exercise accompanied by a negative social evaluation from the experimenter. After completing the stressor task, the subject provided a stress-induced typing sample ("*Typing 2*"), followed by a second 15-minute rest period to return them to a neutral state. A second neutral typing sample was then obtained ("*Typing 3*"). Between each of these phases, subjects had a 2-minute break to fill out the State-Trait Anxiety Inventory (STAI) and NASA Taskload Work Index (NASA-TLX) questionnaires regarding their stress state.

The analysis of the ECG and BP data, as well as the questionnaires, confirm that the subjects experienced stress during the stressor task.

C. Data Selection

In this work, we focus on analyzing the data recorded during the stressor task. Nevertheless, we also compute the

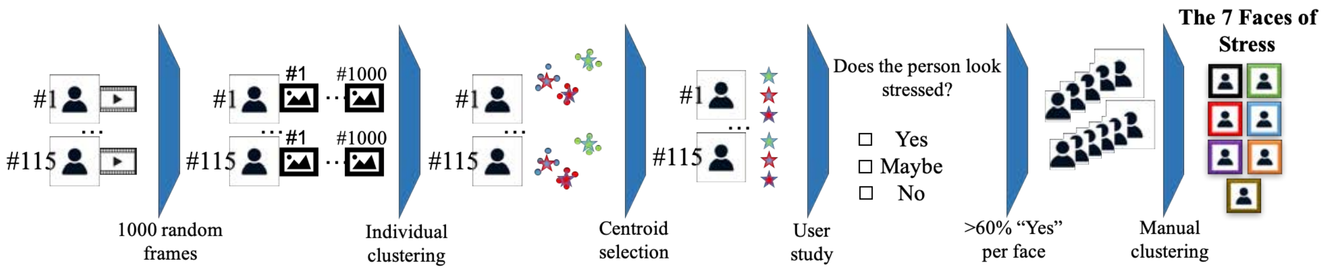

Fig. 4: Overview of our proposed analysis methodology. Our pipeline starts with five-minute videos of 115 subjects. A random selection of 1000 frames per subject follows together with K-means clustering for each individual's data. The centroids (video frames) of each individual's clusters are selected for our user study. Only the centroids with more than 60% "Yes" annotations were selected for manual clustering of the faces. In the end of our pipeline, we obtained seven facial activity patterns of stress.

statistical differences of the AUs intensities during the stress, *Typing 1*, and *Typing 2* phases to confirm that the differences of occurring facial expressions are statistically significant.

Although the stress section of the original experiment was 15 minutes long, for our analysis, we focused on the last five minutes of the stressor task to ensure high stress levels. The complete recording of the *Typing 1* and *2* phases was used which can vary between five to seven minutes depending on the typing speed of the individuals.

D. *Data Visualization*

In our initial examination, we depicted the temporal behavior of each AU per subject. Each plot illustrated distinct AU behaviors, yet specific single AUs were noticeably correlated with the stress phase for each subject. In Fig. 2, for instance, AU 20, identified in [24] as an indicator of fear, exhibited a high correlation with the stress phase in this subject. Another example is AU 2, the outer brow raiser, which was activated during breaks. Similar correlations with various phases were observed in other subjects, confirming the subject-dependent behavior during stress.

Fig. 3 illustrates the results of t-distributed Stochastic Neighbor Embedding (t-SNE) [25] for one subject. The features recorded during different phases, including various typing phases, are color-coded. Notably, the typing phases before and after the stressor are distinctly grouped, while the final typing phase after resting overlaps with the other two.

IV. Methodology

The goal of this work is to define different facial expressions of stress that can be recognized as such by humans. For this purpose, we propose a novel semi-automatic methodology that allows the unveiling of relevant facial activity patterns of stress out of a large amount of data. Fig. 4 shows the overview of our proposed methodology. First, we randomly extracted 1000 frames per subject recording and employed K-means clustering to categorize the data for each individual. The results of the cluster analysis provided us with frequently occurring combinations of AUs during stress for each individual. Next, we used the centroids (video frames) from each individual's clusters for our user study. Only centroids that obtained more than 60% ”Yes” annotations were chosen for manual clustering of the facial expressions. Finally, we identified relevant AUs that co-occur during the different facial stress patterns and the basic emotions.

A. *Feature Extraction*

We extracted 18 distinct AUs using the OpenFace toolbox [3] (see Table I). The intensity values of AUs range from zero to five, with five being the maximum intensity.

To mitigate variations in facial expressiveness among participants, we applied z-normalization to the AUs within each individual. This step was crucial to address the potential skewing effect of highly expressive individuals, who might otherwise disproportionately influence the formation of facial expression clusters. Our objective was not to delineate clusters solely based on different levels of expressiveness (referring to the intensity and frequency of facial responses). Instead, we aimed to extract distinct facial activity patterns that remain consistent across a group of individuals, irrespective of whether the expressions were subtle or more pronounced. The z-normalization within each participant effectively eliminated individual differences in expressiveness, allowing us to focus on the extraction of consistent facial activity patterns. Kunz et al. [20] also employed z-normalization on AUs to cluster pain facial expressions.

For our frequency occurrence analysis, we used binary features of the AUs (Eq. 1)

$$\mathbf{f}_b = (f_{b1} = AU01, f_{b2} = AU02, \cdots, f_{b17} = AU45) \quad \text{with} \begin{cases} f_{bx} = 1, & \text{if } f_{cx} \geq 1 \\ f_{bx} = 0, & \text{otherwise} \end{cases} \tag{1}$$

and for the cluster analysis we used the z-normalized features of continuous AUs (Eq. 2)

$$\mathbf{f}_c = (f_{c1} = AU01, f_{c2} = AU02, \cdots, f_{c17} = AU45) \quad \text{with } f \in [0, 5] \tag{2}$$

TABLE I: Frequency occurrence of facial action units (AUs) during stress, typing 1, and 2 phases. AU 9, AU 10, AU 12, and AU 28 occur during less than 5% of the frames, which is why they are not used for further clustering analysis.

Action Units	**AU Description**	**Stress** f	m	both	**Typing 1** f	m	both	**Typing 2** f	m	both
AU01	Inner brow raiser	13.81	16.37	14.87	13.32	14.32	13.74	13.22	13.67	13.41
AU02	Outer brow raiser	18.1	21.15	19.37	14.28	15.45	14.77	14.93	15.82	15.3
AU04	Brow lowerer	21.19	20.87	21.06	34.45	22.75	29.57	31.07	27.75	29.68
AU05	Upper lid raiser	75.5	80.66	77.65	61.43	69.81	64.93	60.42	67.22	63.26
AU06	Cheek raiser	16.38	24.12	19.61	6.2	16.67	10.57	**4.63**	18.29	10.33
AU07	Lid tightener	41.06	41.37	41.19	35.13	35.86	35.43	32.3	31.95	32.16
AU09	Nose wrinkler	**3.27**	**3.16**	**3.22**	**2.45**	**2.66**	**2.54**	**1.9**	**3.62**	**2.62**
AU10	Upper lip raiser	**3.97**	**4.62**	**4.24**	5.14	**0.97**	**3.4**	5.29	**1.69**	**3.79**
AU12	Lip corner puller	**4.17**	**2.44**	**3.45**	**4.68**	**2.27**	**3.67**	5.41	**1.85**	**3.92**
AU14	Dimpler	6.08	5.35	5.78	8.76	13.48	10.73	9.04	12.46	10.47
AU15	Lip corner depressor	13.19	20.09	16.07	7.69	13.43	10.09	7.5	15.81	10.97
AU17	Chin raiser	27.3	27.38	27.34	17.72	16.87	17.37	16.51	18.84	17.48
AU20	Lip stretcher	12.0	19.05	14.94	7.56	8.85	8.1	7.21	11.16	8.86
AU23	Lip tightener	24.45	33.44	28.2	29.75	43.57	35.52	23.94	44.5	32.52
AU25	Lips part	16.34	14.04	15.38	9.77	9.0	9.45	10.09	9.05	9.65
AU26	Jaw drop	12.05	10.52	11.41	7.45	7.57	7.5	8.14	6.94	7.64
AU28	Lip suck	**0.87**	**1.55**	**1.15**	**0.69**	**0.63**	**0.66**	**0.6**	**0.91**	**0.73**
AU45	Blink	16.24	14.56	15.54	17.86	18.66	18.19	18.52	18.73	18.61

B. Frequency Occurrence of Facial Action Units

To determine the AUs for inclusion in our cluster analysis, we assessed the frequency of each AU during stress. For comparison, we also computed the frequency occurrences (FOs) during the typing 1 and 2 phases. Only AUs that manifested in at least 5% of the recorded stress segments were chosen for additional analyses. Fourteen of the eighteen AUs occurred more often than 5% and are detailed in Table I. We also computed the statistical significance of the differences between the frequency occurrences during stress and the typing phases. Similar AUs showed a p-value < 0.05 when computing the statistical significance of the AUs FOs during stress and typing 1 and stress and typing 2. In both analyses, AU 2, 4, 5, 6, 14, 15, 17, 20, 25, 26 show significantly different occurrences during stress and the typing phases. In the analysis with typing 1, AU 23 also showed a p-value < 0.05, while in the analysis with typing 2, AU 7 and AU 45 were additionally statistically significant.

C. Initial Cluster Analysis

For our initial cluster analysis, we applied the z-normalization on the AUs with FO above 5%. We randomly chose 100 samples per subject from the stress video segment and applied the K-means algorithm using different cluster numbers (K=1 to 21) on a total of 1500 samples. The clustering results were visually not different from each other, which is why we proceeded with performing clustering on data of each individual separately. We selected 1000 samples per subject randomly and also performed K-means clustering with K varying between 1 and 10. We applied the elbow method [32] to choose the ideal cluster number, which for most subjects was three. We also computed the silhouette score [30] for each clustering of the subjects' samples. Fig. 5 shows the clustering results for nine of the 115 subjects and how the silhouette score indicates the visual difference between the clusters. Clusters with low silhouette scores show minimal differences in facial expressions. Clusters with high silhouette scores, on the other hand, show well-distinguishable facial expressions. What is also noticeable, is the variety of facial expressions recognized by performing clustering for each individual. The clusters show unexpressive faces, smiles, pressed lips, open mouths, and asymmetric dimplers.

D. User Study

Given that our initial clusters were automatically created from randomly extracted frames of the stress videos, it is not guaranteed that the obtained clusters show facial expressions that can be recognized as stress. To further refine the data to obtain facial activity patterns that show stress, we performed a user study. Instead of using all 1000 frames per individual, we show the cluster centroids (frames in the center of the clusters) of each subject. Given that the silhouette score of the clusters indicates the visual difference between them, we chose either one frame from one cluster, two clusters, or each cluster to avoid including frames with similar facial expressions. With this selection criteria, we obtained 208 cluster centroids to be shown in our user study.

A total of 17 people (eight female and nine male) took our study online with ages between 25 and 44. In the introduction of the study, we inform that pictures of different people working on a computer will be shown during this study and that they should indicate whether the person looks stressed or not. During the study, individual frames are shown with the question "Does the person look stressed?". The participants can answer with "Yes", "Maybe", or "No". The study was split into four batches to allow breaks in between, showing the 208 frames in a randomized order. Answering each batch took an average of 6 minutes.

In crowd-sourced annotation, the presence of noisy annotations is common, especially when non-expert annotators are involved, as spammers and malicious workers may contribute

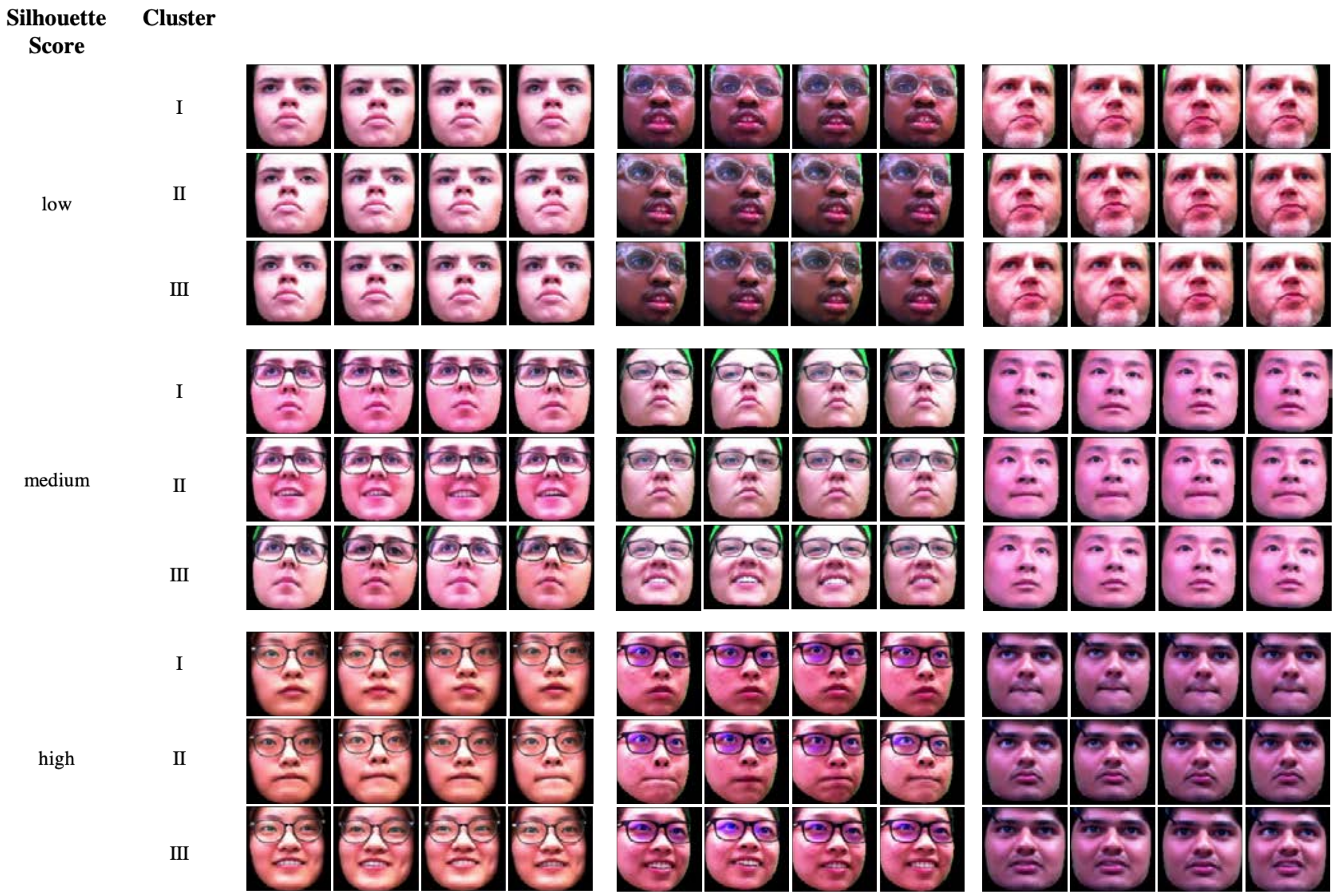

Fig. 5: Examples of clusters obtained by performing clustering on the data of each individual separately. The figure shows the impact of the silhouette score on the differences shown in the clusters. Clusters with low silhouette score (0 to 0.2) show no visible differences between the AUs. Clusters with medium silhouette score (0.2 to 0.4) show two visually distinct clusters and clusters with high silhouette score (> 0.4) show visual differences among all clusters.

[6]. To address this, we assessed the percentage agreement of each annotator's annotations by comparing them to a preliminary majority vote. The analysis revealed that 12 annotators had an agreement lower than 60% with the preliminary majority vote. To ensure the highest annotation quality, we selected the remaining 5 annotators. To validate the high inter-rater agreement, we computed Cohen's kappa [27] pairwise for the 5 annotators, obtaining an average agreement score of 0.42, which is considered moderate.

E. Final Clustering

For final clustering, we selected the data samples from the user study that had a "Yes" annotation for more than 60% of the answers. We applied again K-means clustering on the final data selection, however, only two clusters showed visually consistent similarities over the samples. For this reason, we decided to manually group the samples following common similarities. Two authors of this work performed the manual clustering separately and reached the same clusters. For each cluster, we computed the mean of the z-normalized AU intensities. Following the definitions of which AUs are present during the basic emotions [12], [9], [28] we identified the AUs that co-occur during our defined facial activity patterns of stress and the basic emotions.

V. Results: The Seven Faces of Stress

In the following, we will describe the results obtained through manual clustering of the samples that were identified to show stressed individuals during the user study. Fig. 6 provides an overview of samples belonging to the individual clusters, as well as boxplots of the z-normalized AUs occurring in each cluster. Each cluster shows one very characteristic expression which we used to describe the clusters for easier comprehension.

A. Pressed lips - Cluster 1

This cluster had the most samples in our analysis (24.2 %). The characteristic pressed lips show high intensities for AU 23 (lip tightener), followed by AU 14 (dimpler) and AU 20 (lip stretcher). AU 20 is also a characteristic AU during fear and AU 23 during anger. The focused eyes as well as the pressed lips transmit a tensioned person. This facial activity pattern also can be seen frequently in Fig. 5 in the row with the highest silhouette score.

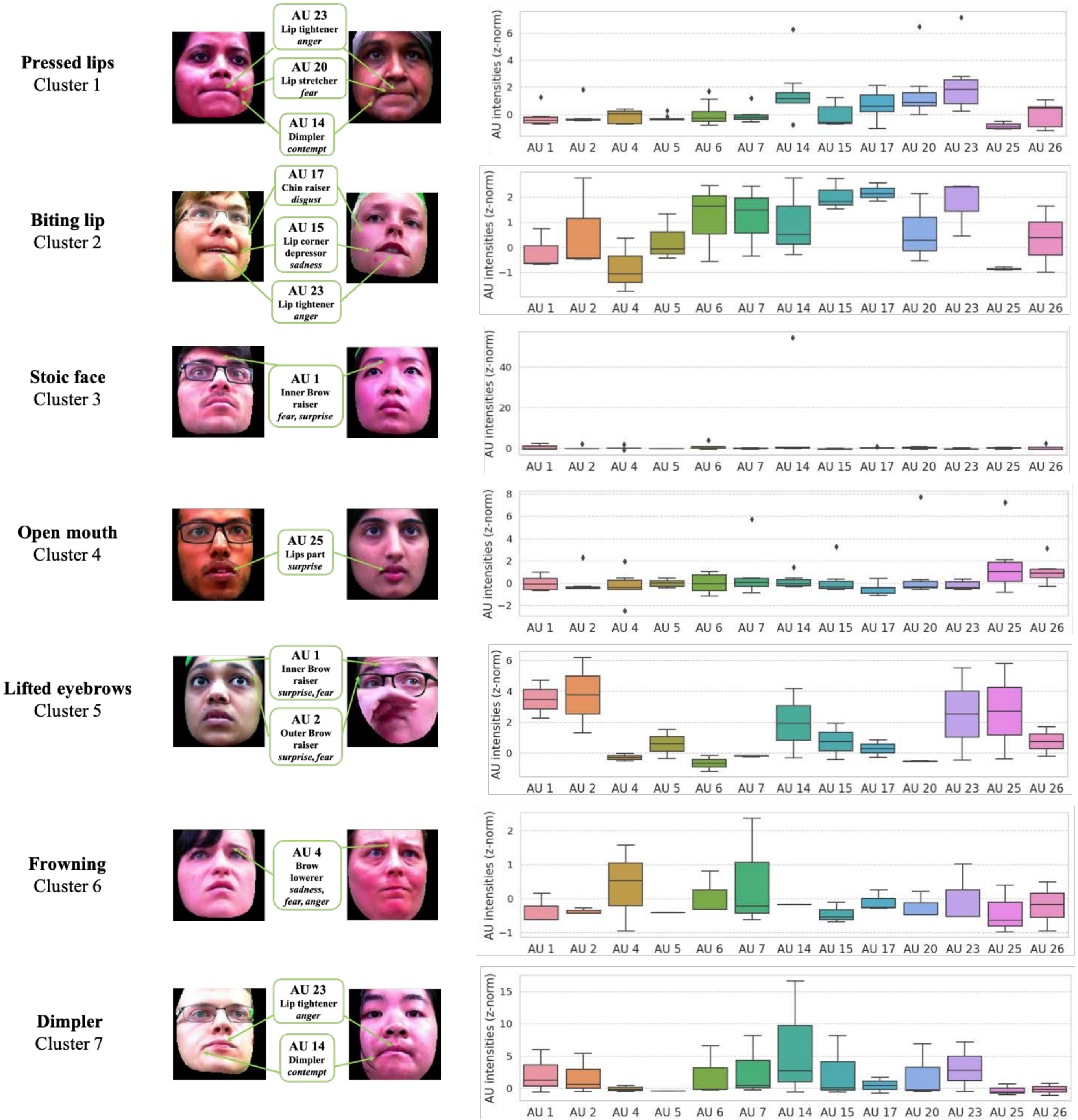

Fig. 6: Results of manual clustering of samples identified as showing a stressed person. On the right of each sample, the mean of the z-normalized AUs intensities is shown with boxplots. The AUs with the highest means are indicated on the samples of the clusters, together with the basic emotions that show the activity of the same AU.

B. Biting lip - Cluster 2

In cluster two, individuals are biting their lower lip. In FACS however, there is a dedicated AU for biting the lip (AU 32), the tool we used cannot detect that AU. The AU with the highest intensity is instead AU 23 (lip tightener) followed by AU 17 (chin raiser) and AU 15 (lip corner depressor). The mentioned AUs are also active during anger, disgust, and sadness respectively.

C. Stoic face - Cluster 3

Cluster three shows a stoic face with mostly low AU intensities. Nevertheless, a slight activity of AU 1 (inner brow raiser) is visible. This facial activity pattern was present in 15% of our final clustering samples. It is also frequent in the individual clustering results in Fig. 5. The inner brow raiser is also characteristic during fear and surprise.

D. Open mouth - Cluster 4

In cluster four individuals have all their mouths open. Although similar to cluster three all AU intensities are low, AU 25 (lips part) shows the highest mean. This cluster was the second most frequent among the samples containing 18% of the samples. Parted lips are also characteristic of the basic emotion of surprise.

E. Lifted Eyebrows - Cluster 5

In cluster five samples show strongly lifted eyebrows. The highest mean values are shown in AU 2 (outer brow raiser) and AU 1 (inner brow raiser). Both AUs are characteristic of the basic emotions of fear and surprise.

F. Frowning - Cluster 6

Cluster six contains samples with predominantly frowned eyebrows. AU 4 (brow lowered) has the highest mean compared to the remaining AUs. Eyebrow frowning is also characteristic during sadness, fear, and anger. In this cluster slight variations of the lower face are visible, showing parted or pressed lips.

G. Dimpler - Cluster 7

The final identified cluster shows a predominantly unilateral dimpler expression which has been associated with the basic emotion of contempt. AU 23 (lip tightener) has a similarly high mean value and is also present in anger.

VI. Discussion

Through our proposed semi-automatic methods we distilled seven facial activity patterns of stress from initially almost 600 minutes of video recording from 115 subjects. Our user study ensured that our proposed facial patterns are recognized as stressed individuals. We also were able to show that the characteristic AUs for each pattern also co-occur during the basic emotions of fear, anger, surprise, sadness, and contempt which indicates that human annotators associate stress with negative emotions.

Although previous work trained successful stress classifiers using solely AUs, it remained unclear how stress is expressed through facial expressions. A major challenge is to find relevant frames that show facial expressions of stress, especially as stress-related facial patterns are not universal. Similar to previous work, we can confirm that components of negative emotions co-occur during our seven facial activity patterns of stress [13], [5]. Compared with previous work that focused on statistical analysis of the occurrence of AUs during stress, we can confirm the result from Blasberg et al. [5] that AU 5 (upper lid raiser) occurred more often during stress than the typing phases, however, it is not one of the main characteristics of any of our proposed seven facial activity patterns of stress.

Although we assumed that stress is shown through different facial activity patterns, we were surprised to clearly distinguish seven facial activity patterns during stress. Our initial clustering of facial expressions of individuals during stress had already shown that some of the subjects in the videos show three very distinct facial patterns during stress. Given that per subject, we chose a maximum of three different facial expressions to be shown to human annotators, it was unexpected to find seven instead of three - more than twice as many - clearly distinguishable facial patterns showing stress. Although our initial clustering results showed that several individuals smiled during the stress activity, human annotators did not relate the smiles with stress. We believe that the occurring smiles are a result of nervousness which has been shown in previous research papers [2], [4]. However, given that the user study was designed using single frames instead of video segments, we hypothesize that the lack of temporal context impeded annotators from recognizing nervousness in the smiles.

VII. Conclusions and Future Work

In this work, we propose a novel semi-automatic method that allows to obtain relevant facial patterns of stress from a large data pool with high variability between individuals. We combined clustering, statistical analysis, and human annotation to obtain seven facial activity patterns of stress by analyzing almost 600 min of video material of 115 different individuals during a stressful task. The resulting seven facial expressions of stress show different characteristic facial expressions and emphasize the variety of facial activity patterns during stressful situations. Although we cannot state that these are the only facial activity patterns showing stress, this work provides the first proposal of stress facial patterns in literature that is not solely based on statistical analysis, but that contains human annotation confirming the recognition of stress. We believe that our work will serve researchers from computer science, human-machine interaction, as well as psychology in advancing systems that allow to improve not only stress detection but also the interaction with stressed individuals. In this work, we did not take into account the temporal sequence of AUs within each activity pattern. Addressing the temporal sequence should be considered the next step to comprehensively capture the facial language of stress.

References

[1] J. Almeida and F. Rodrigues. Facial expression recognition system for stress detection with deep learning. In *ICEIS (1)*, pages 256–263, 2021.

[2] Z. Ambadar, J. F. Cohn, and L. I. Reed. All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of nonverbal behavior*, 33:17–34, 2009.

[3] T. Baltrušaitis, P. Robinson, and L.-P. Morency. Openface: an open source facial behavior analysis toolkit. In *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*, pages 1–10. IEEE, 2016.

[4] J. Belluso. *A Nervous Smile*. Dramatists Play Service Inc, 2006.

[5] J. U. Blasberg, M. Gallistl, M. Degering, F. Baierlein, and V. Engert. You look stressed: A pilot study on facial action unit activity in the context of psychosocial stress. *Comprehensive Psychoneuroendocrinology*, page 100187, 2023.

[6] A. Burmania, S. Parthasarathy, and C. Busso. Increasing the reliability of crowdsourcing evaluations using online quality assessment. *IEEE Transactions on Affective Computing*, 7(4):374–388, 2015.

[7] S. Carolan, P. R. Harris, and K. Cavanagh. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *Journal of medical Internet research*, 19(7):e271, 2017.

[8] D. F. Dinges, R. L. Rider, J. Dorrian, E. L. McGlinchey, N. L. Rogers, Z. Cizman, S. K. Goldenstein, C. Vogler, S. Venkataraman, and D. N. Metaxas. Optical computer recognition of facial expressions associated with stress induced by performance demands. *Aviation, space, and environmental medicine*, 76(6):B172–B182, 2005.

[9] S. Du and A. M. Martinez. Compound facial expressions of emotion: from basic research to clinical applications. *Dialogues in clinical neuroscience*, 17(4):443–455, 2015.

[10] P. Ekman, W. V. Freisen, and S. Ancoli. Facial signs of emotional experience. *Journal of personality and social psychology*, 39(6):1125, 1980.

[11] P. Ekman and W. V. Friesen. *Facial Action Coding System: Investigatoris Guide*. Consulting Psychologists Press, 1978.

[12] W. V. Friesen, P. Ekman, et al. Emfacs-7: Emotional facial action coding system. *Unpublished manuscript, University of California at San Francisco*, 2(36):1, 1983.

[13] H. Gao, A. Yüce, and J.-P. Thiran. Detecting emotional stress from facial expressions for driving safety. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 5961–5965. IEEE, 2014.

[14] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis. Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing*, 2019.

[15] G. Giannakakis, M. R. Koujan, A. Roussos, and K. Marias. Automatic stress analysis from facial videos based on deep facial action units recognition. *Pattern Analysis and Applications*, pages 1–15, 2022.

[16] G. Giannakakis, M. Pediaditis, D. Manousos, E. Kazantzaki, F. Chiarugi, P. G. Simos, K. Marias, and M. Tsiknakis. Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control*, 31:89–101, 2017.

[17] D. Hernando, S. Roca, J. Sancho, Á. Alesanco, and R. Bailón. Validation of the apple watch for heart rate variability measurements during relax and mental stress in healthy subjects. *Sensors*, 18(8):2619, 2018.

[18] S. hon Lau. Stress Detection for Keystroke Dynamics. Technical Report CMU-ML-18-104, Carnegie Mellon University, May 2018.

[19] E. Howe, J. Suh, M. Bin Morshed, D. McDuff, K. Rowan, J. Hernandez, M. I. Abdin, G. Ramos, T. Tran, and M. P. Czerwinski. Design of digital workplace stress-reduction intervention systems: Effects of intervention type and timing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2022.

[20] M. Kunz and S. Lautenbacher. The faces of pain: a cluster analysis of individual differences in facial activity patterns of pain. *European Journal of Pain*, 18(6):813–823, 2014.

[21] R. S. Lazarus. Psychological stress and the coping process. 1966.

[22] R. S. Lazarus and S. Folkman. *Stress, appraisal, and coping*. Springer publishing company, 1984.

[23] J. S. Lerner, R. E. Dahl, A. R. Hariri, and S. E. Taylor. Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses. *Biological psychiatry*, 61(2):253–260, 2007.

[24] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *2010 ieee computer society conference on computer vision and pattern recognition-workshops*, pages 94–101. IEEE, 2010.

[25] L. v. d. Maaten and G. Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.

[26] L. M. Mayo and M. Heilig. In the face of stress: Interpreting individual differences in stress-induced facial expressions. *Neurobiology of stress*, 10:100166, 2019.

[27] M. L. McHugh. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282, 2012.

[28] E. Morales-Vargas, C. Reyes-García, and H. Peregrina-Barreto. On the use of action units and fuzzy explanatory models for facial expression recognition. *PloS one*, 14(10):e0223563, 2019.

[29] M. Norden, O. T. Wolf, L. Lehmann, K. Langer, C. Lippert, and H. Drimalla. Automatic detection of subjective, annotated and physiological stress responses from video data. In *2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8. IEEE, 2022.

[30] K. R. Shahapure and C. Nicholas. Cluster quality analysis using silhouette score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, pages 747–748. IEEE, 2020.

[31] Z. Sun, A. Vedernikov, V.-L. Kykyri, M. Pohjola, M. Nokia, and X. Li. Estimating stress in online meetings by remote physiological signal and behavioral features. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*, pages 216–220, 2022.

[32] M. Syakur, B. Khotimah, E. Rochman, and B. D. Satoto. Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP conference series: materials science and engineering*, volume 336, page 012017. IOP Publishing, 2018.

[33] P. E. Velmovitsky, P. Alencar, S. T. Leatherdale, D. Cowan, and P. P. Morita. Using apple watch ecg data for heart rate variability monitoring and stress prediction: A pilot study. *Frontiers in Digital Health*, 4:1058826, 2022.

[34] P. E. Velmovitsky, M. Lotto, P. Alencar, S. T. Leatherdale, D. Cowan, and P. P. Morita. Can heart rate variability data from the apple watch electrocardiogram quantify stress? *Frontiers in Public Health*, 11, 2023.

[35] C. Viegas, S.-H. Lau, R. Maxion, and A. Hauptmann. Towards independent stress detection: A dependent model using facial action units. In *2018 International Conference on Content-Based Multimedia Indexing (CBMI)*, pages 1–6. IEEE, 2018.

[36] J. Whitehouse, S. J. Milward, M. O. Parker, E. Kavanagh, and B. M. Waller. Signal value of stress behaviour. *Evolution and human behavior*, 43(4):325–333, 2022.