Emotion Prediction - Analysis of emotion-related features

Dr. Carmen Bisogni
University of Salerno
Supervisor

Dr. Marcantuono Veronica
University of Salerno
Computer Science Department
0522501198

Dr. Saporito Antonio
University of Salerno
Computer Science Department
0522501135

Dr. Sorrentino Andrea
University of Salerno
Computer Science Department
0522501183

Abstract

The project description can be summarised in these steps:

(1) *Using MediaPipe extract the 468 landmarks from the face;*
(2) *Using MediaPipe, which also allows the selection of specific landmarks, analyse those of the periocular area (Left and Right eye together), of the mouth, of the eyebrows (Left and Right eyebrows together);*
(3) *For each frame create csv containing the extracted landmarks (there are 3 coordinates so they will be 468x3 matrices), thus creating 3 csv per frame;*
(4) *For each video sequence (provided in the CK+ Dataset in the form of frames), analyse the resulting csv to save two types of distances (Local and Global) of the 468 landmarks (Euclidean Distance). For each video sequence two csv of the distances, where each row represents a frame and each column a distance (local in one csv, global in the other). This must be repeated for all 3 zones;*
(5) *Analysing the trend of distances (Local and Global) to highlight the point at which micro becomes macro. Repeat this operation for each emotion found, highlighting the differences between emotions, and between the same emotion but different subjects;*
(6) *Carry out the calculation of averages;*
(7) *Calculate cosine similarity and Pearson's correlation for subject and emotion;*
(8) *Show tables of results;*

1 INTRODUCTION

The project in question, directed by Dr. Carmen Bisogni, aimed to carry out a statistical analysis on the detection of micro and macro expressions. In an initial phase, a survey of the 468 landmarks, i.e. the points inherent to the face, was carried out by means of MediaPipe, which also allows specific landmarks to be selected, the specific points of interest were analysed, those of the periocular area (Left and Right eye together), the mouth and the eyebrows (Left and Right eyebrows together). The project was divided into three sections, each managed by a specific member, Dr. Marcantuono dealt with the eyes, Dr. Saporito dealt with the lips and Dr. Sorrentino dealt with the eyebrows. The work between the various components, despite being divided, was completed simultaneously through close collaboration. Subsequently, csv files containing the extracted landmarks were created for each frame. The extracted coordinates are 3 so for the face there are a total of 468x3 matrices, but of these points the 10 points for the lips, 16 points per eye, 32 in total, and the 10 points for eyebrows, 20 in total, in the whole the project we examine 62 points.

Once the extraction of the csv was made, we calculate the local and global distances using various techniques, one of which is the Euclidean distance:

(1) *A local distance is the calculation that is performed, on a single subject and for each subject in the dataset, for example with the Euclidean distance where we compare the previous csv of a frame with the next csv, considering the i-th as the first csv of a frame we will have the comparison between the i-th and the i+1-th, between the i+1-th and the i+2-th, between the i+2-th and the i+3-th etc. up to the comparison between the n-1-th and the n-th csv;*
(2) *A global distance is the calculation that is performed, on a single subject and for each subject in the dataset, for example with the Euclidean distance where we compare the first csv of a frame with all the other csv's;*

This distance calculation was repeated for all three zones examined. Then for all three zones, the average for each distance was calculated, both local and global, the average is used in conjunction with a similarity index, such as the Cosine similarity metric, that finds the normalised dot product of the two attributes or Pearson Correlation what is a measure of an association between variables. *Specifically we have:*

(1) *Cosine Similarity per Subject and Pearson Correlation per Subject: Where emotions vary according to the subject, the expected result is to have low similarity and low correlation, since the project aims at recognising the emotion and not the subject.*
(2) *Cosine Similarity for Emotion and Pearson Correlation for Emotion: Where subjects are made to vary according to emotion, the expected result is to have high similarity and to have high correlation, since the project aims at emotion recognition.*

2 RELATED WORKS

As related work, Google Scholar examined the project progetto "MediaPipe Hands: On-device Real-time Hand Tracking" [1], where the authors explain that they presented a real-time hand tracking pipeline on the device that predicts the hand skeleton from a single RGB camera for AR/VR applications. Where the pipeline consists of two models, a palm detector and a hand reference model. It is implemented via MediaPipe, a framework for creating cross-platform ML solutions. The proposed model and pipeline architecture demonstrates real-time inference speed on mobile GPUs and high prediction quality.

Or the project: "MediaPipe: A Framework for Building Perception Pipelines" [2], where the authors explain that building applications that perceive the surrounding world is challenging and a developer has to perform a number of choices and operations inherent to development. A developer can use MediaPipe to build prototypes by combining existing perception components, to progress them to cross-platform applications, and to measure system performance and resource consumption on target platforms, demonstrating that these features allow the developer to focus on algorithm or model development and use MediaPipe as an environment to iteratively improve their application with reproducible results on different devices and platforms.

Or the project: "Real-Time Human Pose Detection and Recognition Using MediaPipe" [3], where the authors explain that the importance of human action recognition has increased disproportionately due to its large-scale application in the field of public safety, games, etc., proposing a framework that detects human action under different conditions and viewing angles that allow the identification of divergent patterns based on different spatio-temporal trajectories. This is made possible through the use of MediaPipe Holistic, which provides pose, face and hand point detection models, which analyses the frames obtained through real-time device feeding using OpenCV through the MediaPipe Holistic model and providing a total of 501 landmarks that are exported as coordinates in a CSV file on which to train a customised multi-class classification model to understand the relationship between the class and the coordinates to classify and detect the customised body language pose.

Or the project: "A Peek at Peak Emotion Recognition" [4], where the authors explain that in the field of facial expression recognition, little attention has been paid to the recognition of peaks of emotion. Aviezer et al. showed that humans have difficulty discerning between positive and negative peaks of emotion. Through behaviour analysis, it was found that despite using very small datasets, features extracted by deep learning models can perform significantly better than humans and that deep learning models, even if trained only on datasets labelled by humans, still outperform humans in this task.

3 PROPOSED SYSTEM

3.1 Dataset analysis and normalisation

Based on the dataset of labels containing the subject's emotion and after an analysis of the content of our dataset, we have:

- Renamed the contents of the dataset (e.g. S005_001, where S005 represents the subject, 001 represents the emotion of the subject). To do this, we used the rename.py code, which allowed us to change the name of all the frames in a given folder;
- Removed folders containing frames of emotions not recognised by the dataset (example: emotions "similar" to the neutral one but with a different expression);

3.2 Label file generation

We have created a code to extract from the dataset all the subjects and the emotions, based on the name of the file. The output will be a csv file called label.csv that will contain all the subjects and their emotions.

3.3 Landmark extraction

To extract the landmarks, we used the functions (for lips, eyes and eyebrow) in the code mediap_util.py

(1) *Problem: In addition to the points of interest, the functions also returned the number of faces, this generated an execution error when calling the function as the number of faces is not necessary for our work;*
(2) *Solution: The return of the number of faces was removed from the function;*

We then created the code csv_landmark.py for each part of the face of interest where the specific function is called. Each function extracts a different number of points:

- *Lips: Extracts 10 points per frame;*
- *Eyes: Is divided into 2 functions, each of which extracts 16 points per eye, for a total of 32 points per frame;*
- *EyeB: is divided into 2 functions, each of which extracts 10 points per eyebrow, for a total of 20 points per frame;*

In order to extract the landmarks, we scanned the entire dataset and generated one csv per image. The landmark extraction was divided into 91 subjects for the train, i.e. approximately 80% of the total number of subjects, and 32 subjects for the test, the remaining approximately 20%.

(1) *Problem: When performing the extraction of landmarks on the entire dataset, we realised that the operation was not successful because after a certain number of processed files, the execution was interrupted due to "RAM memory saturation" (the computer used has 32GB of memory and was processing about 4 thousand files out of about 10 thousand in the database);*
(2) *Solution: We divided the dataset into 9 subfolders containing about 13 Subjects per folder for a total of about 1200 files to be processed. The extraction of the landmarks is done manually by varying the path because if we go through all the sub-folders in one run anyway, we go into memory saturation error;*

3.4 Local and global distances

Using Euclidean distance, we calculated the local distances and global distances between files relating to the same subject. Consequently, for local distances we compared for all subjects the i-th csv file containing the subject's points with the next one, thus generating a csv file for each subject, containing the local distances obtained. For global distances, on the other hand, we compared for all subjects the first csv file of each subject with the i-th csv files of the same subject, thus generating one csv file for each subject, containing the global distances obtained.

3.5 Average calculation

Initially, with the calculation of distances, a single csv file was created containing the distances of all subjects, then we decided to separate the distance file by subject in order to simplify the process of calculating the average, as explained in the solution to the first problem:

(1) *Problem: Having a single csv with all subjects and all distances presented the problem of separating the subjects to be averaged in the csv and consequently not being able to calculate averages from 20% to 60% of the columns referring to the subject;*
(2) *Solution: We decided to separate the distances for each subject so that we would have one csv of the distances for each subject (for both global and local distances). This allowed us to avoid carrying forward three indices, (Subject Start Index; Subject End Index; Subject Number of Columns). By separating the csv's, this issue was solved by averaging 20% to 60% per cent of the subject columns over the entire csv under consideration;*

(1) *Problem: When calculating the averages from 20% to 60% of the columns per subject, we have lists of varying size. Due to this variability, it was not possible to create a csv file containing the averages as lists of different sizes (it generated the index error);*
(2) *Solution: The first solution was to normalise the lists by making them the same size by inserting a NaN value in the empty indexes of the shorter lists. This would have avoided the index out of range error but would have inserted dummy values into the file. Next, in order to avoid falsifying the data, it was decided to select the middle column of the csv and (After averaging the size of the lists resulting from the 20% - 60% calculation) take the 4 columns before and 4 after the middle column. As a result, the index out of value error was overcome by having all lists of the same size;*

(1) *Problem: In taking the middle 8 columns, we realised another problem, namely that not all subjects have a sufficient number of columns to be extracted. In fact, some distance rows have less than 8 columns and others more than 30;*
(2) *Solution: In the calculation of the averages, an additional control has been inserted allowing files with less than 8 columns to be skipped. The end result will be to have 2 csv per part of the face considered, one with averages of local distances and one with averages of global distances;*

N.B: Calculating averages in this way makes the subsequent operations we will perform on the average files inaccurate, since the average performed on a file with 10 columns will be very different from a file with 30 columns (even if referring to the same emotion). This is because in a file with 10 columns the change in expression is much more evident than in a file with 30/35 columns, where perhaps the change occurs much more slowly.

3.6 Cosine similarity for subject and emotion

Cosine similarity is a heuristic technique for measuring the similarity between two vectors by calculating the cosine between them.

To calculate similarity, we decided to first fix the subject by varying the emotion and then fix the emotion and vary the subject. We expect to obtain a similarity:

- Low when, having fixed the subject, we vary the emotion;
- High when, having fixed the emotion, we make the subject vary;

3.6.1 ***Cosine similarity by subject:*** We passed in the input files of averages for local distances and for global distances. What is done in the subject similarity code is basically to create a dictionary containing all the subjects with their emotions and data. Consequently we will have a dictionary of dictionaries, where we will specifically have:

- Subject;
- emotions;
- Subject_emotion values (Taken from the averages file);

The outer dictionary will contain all subjects, the inner dictionaries will contain subject emotions and emotion data. With the function: 1 - spatial.distance.cosine() of the scipy library we are going to perform similarity on 2 lists, consequently to perform the similarity calculation by fixing the subject, we are going to scroll through all the subject dictionaries, and as lists we pass the emotion ones 2 by 2. In particular, the code will compare the emotion taken into consideration for the subject with all the other emotions, consequently the final result for a subject that has 7 emotions, will be a 6*7 matrix, structured as follows:

- Column 1: Will contain 7 values resulting from the comparison of emotion 1 with itself, and all subsequent ones;
- Column 2: Will contain 7 values resulting from the comparison of emotion 2 with the previous one, with itself and all subsequent ones;
- Column 3: Will contain 7 values resulting from the comparison of emotion 3 with the previous ones, with itself and all subsequent ones;
- Etc.

(1) *Problem: Not all subjects have the same number of emotions so the index out of range error would occur again;*
(2) *Solution: To avoid the index out of range error again, we decided to normalise the lists by inserting 0.5 in the cells that should contain the similarity result with the other emotions;*

The output will be 2 csv files, one for running the similarity applied on the local distance averages file and one on the global distance averages file. These csv files will contain the similarities for all subjects.

3.6.2 ***Cosine similarity by emotion:*** We passed as input the average files for local and global distances. Like the similarities by subject, here instead we have a dictionary containing the dictionaries of all emotions and their values.

(1) *Problem: Not all emotions have the same number of subjects so the index out of range error would occur again;*
(2) *Solution: To avoid the index out of range error again, we decided to normalise the lists by inserting 0.5 in the cells that should contain the similarity result with other subjects;*

3.7 Pearson correlation by subject and emotion

Correlation [5] in the broadest sense is a measure of an association between variables. In correlated data, the change in the magnitude of 1 variable is associated with a change in the magnitude of another variable, either in the same (positive correlation) or in the opposite (negative correlation) direction. Most often, the term correlation is used in the context of a linear relationship between 2 continuous variables and expressed as Pearson product-moment correlation. The Pearson correlation coefficient is typically used for jointly normally distributed data (data that follow a bivariate normal distribution). For nonnormally distributed continuous data, for ordinal data, or for data with relevant outliers, a Spearman rank correlation can be used as a measure of a monotonic association. Both correlation coefficients are scaled such that they range from -1 to 1, where 0 indicates that there is no linear or monotonic association, and the relationship gets stronger and ultimately approaches a straight line (Pearson correlation) or a constantly increasing or decreasing curve (Spearman correlation) as the coefficient approaches an absolute value of 1. Hypothesis tests and confidence intervals can be used to address the statistical significance of the results and to estimate the strength of the relationship in the population from which the data were sampled.

To calculate Pearson correlation, as for similarity we decided to first fix the subject by varying the emotion and then fix the emotion and vary the subject.

3.7.1 ***Pearson correlation by subject:*** We passed as input the averages files for local distances and for global distances. The Pandas library allows us to perform the Pearson correlation calculation with the function corr(method="pearson") The result of this function is an n*n table containing the values of the comparisons made where we will have the same header for both rows and columns.

	S128_002	S128_003	S128_005
S128_002	1	0.9996442962	0.9996183088
S128_003	0.9996442962	1	0.9991267875
S128_005	0.9996183088	0.9991267875	1

What has been done then is to go and select the subjects one at a time, perform the correlation calculation.

(1) *Problem: Given the matrix resulting from the correlation, it is complicated to enter the correlations of all subjects in a single csv file;*
(2) *Solution: We decided to create a csv for each correlation calculated;*

3.7.2 ***Pearson correlation by emotion:*** We passed in input the averages files for local distances and for global distances. Basically, the code works in the same way as the correlation by subject, with the difference that in order to go fix the emotion, we fetch the column sorting from the similarity by emotion file.

3.8 Table Generation

As a final step, we implemented tables in order to show the resulting data in the previously generated files on the screen. A switchcase was inserted that first allows you to choose which file to open and then choose the subject whose data you want to show. To select a subject whose data is to be shown, the label.csv file containing all subjects and emotions must be consulted.

(1) *Problem: Not all csv files can be shown on the screen, as some tables would be too large and the data very small;*
(2) *Solution: We have decided to show in the tables only those files that make it possible to generate tables that are not excessively large. The following files will then be displayed:*

```
1. Lips_similarity_Sogg_LocalAverage.csv
2. Lips_similarity_Sogg_GlobalAverage.csv
3. Lips_average_locali.csv
4. Lips_average_globali.csv
5. Eyes_similarity_Sogg_LocalAverage.csv
6. Eyes_similarity_Sogg_GlobalAverage.csv
7. Eyes_average_locali.csv
8. Eyes_average_globali.csv
9. EyeB_similarity_Sogg_LocalAverage.csv
10. EyeB_similarity_Sogg_GlobalAverage.csv
11. EyeB_average_locali.csv
12. EyeB_average_globali.csv
```

4 DATASET COHN - KANADE [CK+]

The CK+ dataset [6] consists of 593 video sequences from 123 participants. Each sequence contains images beginning from onset (neutral frame) and progressing to the peak expression (last frame). The label associated with each sequence is depicted from the peak expression. The dataset contains images for seven different expressions: anger, contempt, fear, disgust, happiness, surprise, and sadness. The images have a resolution of 640x480 pixels.

4.1 Image Data [7]

Facial behavior of 210 adults was recorded using two hardware synchronized Panasonic AG-7500 cameras. Participants were 18 to 50 years of age, 69 % female, 81%, Euro-American, 13% Afro-American, and 6% other groups. Participants were instructed by an experimenter to perform a series of 23 facial displays; these included single action units and combinations of action units. Each display began and ended in a neutral face, any exceptions noted. Image sequences for frontal views and 30-degree views were digitized into either 640×480 or 640×480 pixel arrays with 8- bit gray-scale or 24-bit color values.

5 WORK FLOW

As specified in the section on calculating averages, the data obtained may not be very accurate as the subjects in the Dataset have varying numbers of frames. Therefore, the averages obtained could make the expression change data very different from each other as it is possible for a very noticeable expression change to occur in a subject with 10 frames rather than one with 30 frames. For example, for subject 126 emotion 1, the change of expression on the face occurs very gradually, and it is not possible to perceive this change considering only the 8 central columns. In contrast, for emotion 7, the change of expression is much more rapid and visible in fewer frames.

5.1 Analysis of cosine similarity results by subject

Analysing the results obtained per section, with respect to the variation of emotion while staring at the subject, we noticed that the variation is minimal, since taking into consideration the data of the single sections (eyes, eyebrows and mouth) they are insufficient to detect a real change. In fact, as can be seen from the table below, taking subject 129 into consideration, it can be seen that minimal change was found when varying the emotion. The table below shows the variation data for the eye points, compared to the file of averages calculated on local distances.

S129_001	S129_003	S129_004	S129_005	S129_007
1	0.9605318601	0.9509587215	0.9938913637	0.9937984567
0.9605318601	1	0.9994666586	0.9853436462	0.9854922805
0.9509587215	0.9994666586	1	0.9792583802	0.9794366642
0.9938913637	0.9853436462	0.9792583802	1	0.9999931705
0.9937984567	0.9854922805	0.9794366642	0.9999931705	1

Recall that the first column contains the similarity of the first emotion of subject 129 with the other emotions of subject 129. For a clearer reading, the first value in the first column refers to the comparison of S129_001 with itself, the second value refers to S129_001 with S129_003, etc. For a better evaluation of the results obtained, images of the subject are also shown below:

5.2 Analysis of cosine similarity results by emotion

Analysing the results obtained per section, compared to the variation of the subject by fixing the emotion, we noticed that even here the variation is minimal, as the data obtained refer only to the section taken into consideration. Consequently, even when varying the subject, it is difficult to say that there is actual similarity between the same emotion for different subjects. The table below shows the variation data for the eye points, compared to the file of averages calculated on local distances.

S126_001	S127_001	S129_001	S130_001	S133_001	S134_001
1	0.9919703429	0.9177427058	0.9977530483	0.9976530852	0.956122994
0.9919703429	1	0.9605971813	0.9982078352	0.9809866938	0.9854871707
0.9177427058	0.9605971813	1	0.9422739356	0.888406168	0.9938158029
0.9977530483	0.9982078352	0.9422739356	1	0.9908331402	0.9735932493
0.9976530852	0.9809866938	0.888406168	0.9908331402	1	0.9338344192
0.956122994	0.9854871707	0.9938158029	0.9735932493	0.9338344192	1

For a better evaluation of the results obtained, images of the subject are also shown below:

The same applies to the other facial sections taken into account. Even on global distances, the resulting table shows data with minimal variation.

5.3 Pearson correlation result analysis by subject

When analysing the results obtained with Pearson correlation by staring at the subject and varying the emotion, we found that these were much closer to 1 (Understood as a percentage, therefore 1 indicates 100%), precisely because of the discourse referred to the averages. The table below shows the variation data for the eye points, compared to the averages calculated on the local distances.

	S129_001	S129_003	S129_004	S129_005	S129_007
S129_001	1	0.999974447	0.9999643324	0.9999502105	0.9999646625
S129_003	0.999974447	1	0.9999670724	0.9999700312	0.9999727564
S129_004	0.9999643324	0.9999670724	1	0.9999197175	0.9999377089
S129_005	0.9999502105	0.9999700312	0.9999197175	1	0.9999680407
S129_007	0.9999646625	0.9999727564	0.9999377089	0.9999680407	1

N.B: The same applies to the other sections of the face taken in consideration and also to the table on global distances.

5.4 Pearson correlation result analysis by emotion

For the analysis of the results with respect to Pearson's correlation, by fixing the emotion and varying the subject, we again found the same results. The table below shows the data for the variation for the eye points, compared to the file of averages calculated on local distances.

	S126_001	S127_001	S129_001	S130_001	S133_001
S126_001	1	0.9999806569	0.9999822783	0.9999367365	0.9998547665
S127_001	0.9999806569	1	0.9999591398	0.9999228002	0.9998520504
S129_001	0.9999822783	0.9999591398	1	0.9999530284	0.9998952249
S130_001	0.9999367365	0.9999228002	0.9999530284	1	0.9999766673
S133_001	0.9998547665	0.9998520504	0.9998952249	0.9999766673	1

N.B: The same applies to the other sections of the face taken in consideration and also to the table on global distances.

6 TECHNICAL SPECIFICATIONS

With regard to the technical specifications of the group's terminals used for the development of the project, we can first of all say that Antonio's PC was the main terminal as, of the various terminals, it is the most performing, given that with Andrea's PC and Veronica's

PC it was possible to compute at most 2000 files of the dataset for the various execution operations, while Antonio's PC reached a peak of 4000 computed files, obviously making the execution times much shorter, since with the other terminals, for fairly large blocks of files, memory saturation would occur.

Below are all the specifications for the various terminals:

Device Name:	Antonio's PC
Processor:	11th Gen Intel(R) Core (TM) i7-1165G7 @ 2.80GHz 2.80 GHz
Ram Installed:	32,0 GB DDR4 2400 MHz
Graphics Processing Unit:	NVIDIA GeForce GTX 1650 GDDR5 - 4096 MB
System type:	Windows 11 64-bit operating system
Solid-state drive (SSD):	Samsung 970 Evo Plus M.2 NVMe PCle 500 Gb (3500 MB/R – 3200 MB/W)

Device Name:	Andrea's PC
Processor:	8th Gen Intel(R) Core (TM) i7-8750H @ 2.20GHz 2.20 GHz
Ram Installed:	16,0 GB DDR4 2666 MHz
Graphics Processing Unit:	NVIDIA GeForce GTX 1060 GDDR5 - 6144 MB
System type:	Windows 11 64-bit operating system
Solid-state drive (SSD) 1:	Samsung 970 Evo Plus M.2 NVMe PCle 500 Gb (3500 MB/R – 3200 MB/W)
Solid-state drive (SSD) 2:	Adata SX6000 PCIe M.2 2280 512 GB (1000 MB/R, 800 MB/W)

Device Name:	Veronica's PC
Processor:	11th Gen Intel(R) Core (TM) i7-1165G7 @ 2.80GHz 2.80 GHz
Ram Installed:	16,0 GB DDR4 1600 MHz
Graphics Processing Unit:	NVIDIA GeForce MX450 GDDR5 - 2048 MB
System type:	Windows 11 64-bit operating system
Solid-state drive (SSD):	Samsung 970 Evo Plus M.2 NVMe PCle 500 Gb (3500 MB/R – 3200 MB/W)

7 CONCLUSIONS

In conclusion, we can say that this type of approach to recognise a change from micro- to macro-expression is inefficient, as the data are very similar even when specific subjects and emotions vary. Consequently, in order to better evaluate this approach, one would have to:

(1) *Normalise the database so that each subject has the same number of frames per emotion, so to have more precise averages to work with;*
(2) *Take several sections of the face into account at the same time by considering, for example, the contour of the face or consider points of the entire face;*

REFERENCES

[1] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. 2020. Mediapipe hands: on-device real-time hand tracking. *arXiv preprint arXiv:2006.10214*.

[2] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, et al. 2019. Mediapipe: a framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*.

[3] Amritanshu Kumar Singh, Vedant Arvind Kumbhare, and K Arthi. 2021. Real-time human pose detection and recognition using mediapipe. In *International Conference on Soft Computing and Signal Processing*. Springer, 145–154.

[4] Tzvi Michelson, Hillel Aviezer, and Shmuel Peleg. 2022. A peek at peak emotion recognition. *arXiv preprint arXiv:2205.09791*.

[5] Patrick Schober, Christa Boer, and Lothar A Schwarte. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126, 5, 1763–1768.

[6] Patrick Lucey, Jeffrey F Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. 2010. 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*. IEEE, 94–101.

[7] Ali I Siam, Naglaa F Soliman, Abeer D Algarni, Abd El-Samie, E Fathi, and Ahmed Sedik. 2022. Deploying machine learning techniques for human emotion detection. *Computational Intelligence and Neuroscience*, 2022.