EMO: Intelligent framework for utilizing facial emotions for medical purposes

Human Interaction Design for AI

Damir Mubarakshin
Innopolis University
d.mubarakshin@innopolis.ru

Denis Rangulov
Innopolis University
d.rangulov@innopolis.ru

Alina Chasova
Innopolis University
a.chasova@innopolis.ru

Anastasia Bilskaya
Innopolis University
a.bilskaya@innopolis.ru

ABSTRACT

Obvious connection exist between human health and human emotion. Recognition of this emotions open a new possibilities in medical area. Doctors can extract only valid symptoms from patients. Also, it allows assessing the degree of pain and its impact on the patient's standard of living.

We propose a project called "EMO" which represents system using deep neural networks for utilizing facial emotions of patients from videos produced by usual cameras like phone, tablet or surveillance cameras. It can evaluate how positive and brightly expressed emotions of patients are in different scenarios. Besides the system can be used to assess feelings of medical personal to interrupt them from different working activities if they are tired or emotionally unstable.

1 INTRODUCTION

For several decades, emotion recognition has remained one of the most important problems in the field of human computer interaction. A large portion of the community has focused on categorical models which try to group emotions into discrete categories. The most famous categories are the six basic emotions originally proposed by Paul Ekman [1]: anger, disgust, fear, happiness, sadness, and surprise. These emotions were selected because they were all perceived similarly regardless of culture.

Several datasets have been constructed to evaluate automatic emotion recognition systems such as the extended Cohn-Kanade (CK+) dataset and the Jaffe dataset. In the last few years, several methods based on hand-crafted and, later, learned features have performed quite well in recognizing the six basic emotions. Unfortunately, these six basic emotions do not cover the full range of emotions that a person can express.

An alternative way to model the space of possible emotions is to use a dimensional approach where a person's emotions can be described using a low-dimensional signal (typically 2 or 3 dimensions). The most common dimensions are (i) arousal and (ii) valence. Arousal measures how engaged or apathetic a subject appears while valence measures how positive or negative a subject appears. Dimensional approaches have two advantages over categorical approaches. The first being that dimensional approaches can describe a larger set of emotions. Specifically, the arousal and valence scores define a two dimensional plane 1 while the six basic emotions are represented as points in said plane. The second advantage is dimensional approaches can output time-continuous labels which allows for more realistic modeling of emotion over time. This could be particularly useful for representing video data.

From emotional representation it is possible to retrieve the pattern of changing emotions in a way corresponding to lie.

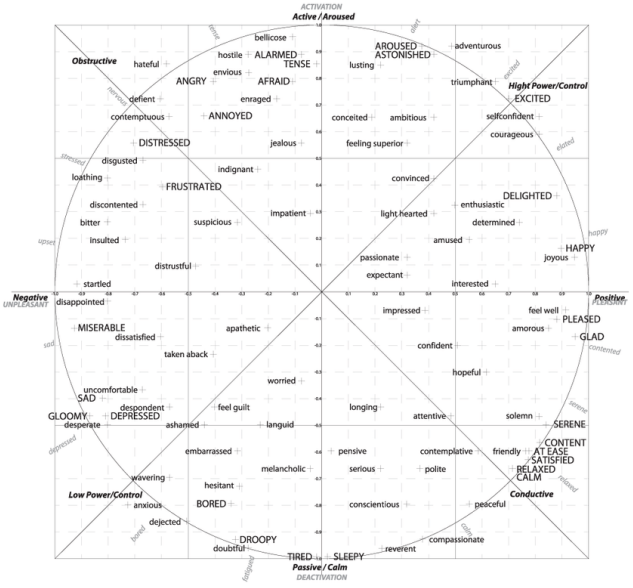

Figure 1: The 2D Emotion Wheel

1.1 Why people lie?

Lying is an integral part of human being. People lie everyday, starting with small doses of lie and ending with larger ones. It is often hard to detect if somebody says the truth or not, but sometimes you must know the verity to save somebody's life or just do not make harm. People resort to lying to each other for so many different reasons that it'd be impossible to list them all. However, one of the most common motives for telling lies is avoiding punishment. This is the primary motivator for both children and adults. [2]

As it was mentioned below, people may lie for many reasons: concealing reward or benefit, protecting someone from harm, self-protection, maintaining privacy, avoiding embarrassment, being polite [2]. These are the most usual and understandable reasons

to explain why people lying. More completely the list of possible causes are provided in [3]. Let's take a look for some of rarely met ones with small descriptions:

(1) Preservation of autonomy - being independent on the situations and people, relieve oneself of responsibility.
(2) Aggression or power - feeling oneself more powerful in the eyes of the somebody, for example to impress the person.
(3) Wish fulfillment - lying to people with the aim to make them feel happy or relax.
(4) Furtherance of self-deception or repression of conflict - lying with the aim to make somebody feel relax on the stressful situation.
(5) Manipulation of others - lying with the aim to get needed utilities.
(6) Assert one's sense of self - often people lie themselves with the aim of self-affirmation.
(7) Solve role conflicts - solving person's inner role conflicts, when competing demands are made on an individual in the fulfillment of his or her multiple social roles.

People's lie is based on protecting themselves from external threats. In spite of this the lie may cause a threat by itself. For example, when it comes to health.

1.2 Why people lie to doctors?

In the interview of Elanie K. Howley with the Angela Fagerlin, professor and chair of population health sciences at University of Utah Health, Fagerlin says that patients often lying because of their life situation. As an example, she told often happening situation, when patient already is taking four medications to manage several conditions and is being prescribed the fifth one. Patient just cannot afford such big expenses or just does not want to hassle. But this is just a small example of the causes of lies.

Thus, life insurance company TermLife2Go collected responses from 500 people and, the results provide some information of the average doctor-patient relationship. They reveal an interesting breakdown between genders: 50% of men say they have lied about alcohol, compared to 32% of women. However, more women (33%) lied about sexual partners than men (21%). The most popular reasons for lies to the doctors were the following:

(1) to avoid discrimination (31%)
(2) to be taken more seriously (22%)

As an individual example one man said he lied to his doctor about alcohol consumption to avoid a lecture about drinking too much.

Usually there exists a lot of reasons why people lie to their doctors, starting from the personal reasons and ending with the physiological problems. Some people lie, in order to minimize or omit information because they don't trust the doctor or the healthcare system. Another reason people might be untruthful is to "get what they want". For example, emphasizing certain symptoms and downplaying others in order to get a certain test or prescription they feel would be beneficial.

The human tendency to give answers we think will reflect well on us rather than the unvarnished truth. Exercise and smoking are prime examples of topics where social desirability bias can have an affect.

1.3 Dangerous Consequences

Lying to the doctors could lead to some very dangerous consequences, such as incorrect diagnoses and, accordingly, incorrect treatment. Lack of open communication between patients and health care providers can affect a person's health. With wrong or unclear information a healthcare provider may miss an important signs or treat the patient with the wrong medicine.

1.4 Lie Detection Methods

Different lie detection methods are evaluated in [4]. We will examine main ones.

Some older methods was based on the belief that God would not let a righteous man suffer and injustice prevail. One them is the hot water test, the accused was ordered to place the hand into a cauldron with boiling water and hold it there for a specified time. If the hand in boiling water showed no traces of scalding or small blisters, it represented a sign of the accused person's claim to be true.

In the field of criminology, there is the graphology method was deemed an appropriate means of verifying the authenticity of documents and signatures. However, graphology was not acknowledged as an appropriate tool for lie detection. Nowadays, this method is used in various areas such as employment profiling (to do a personality profile) or psychological analysis.

The polygraph lie detection device was developed to measure changes in the accused person's blood pressure which were recorded on a graph or chart. The control questions are related to the suspect's crime investigation are asked, but not specifically to the crime. This test measures the suspect's detailed knowledge of a crime that he or she does not want to share. For example, the polygraph examiner might discuss with a suspect several different types of cars, one of which was actually used in committing the crime.

Observation of nonverbal expressions and Voice Stress Analysis. The desire to detect lies is not reflected only in the use of various technical equipment. Observation and attention focused on some specific behavioral expressions have also played important roles. Darwin described Duchenne's work from 1862 which supposes the possibility of revealing the truth by observation of facial expressions. A smile which is the result of experiencing happiness is manifested by constriction of zygomatic major muscle (musculus zygomaticus major) causing the corners of the mouth to lift. In case of electrical stimulation of this muscle, the smile appears to be unnatural. Similarly, this applies to the circular muscles in the eye (orbicularis oculi) which, when constricted, pull the face slightly higher and depress the eyebrows. These two muscles can reveal the true emotional state since their activity can be purposely controlled only with great difficulties.

In addition, Ekman (1996) presented six interpretations as to why we are not successful in lie detection. One of the reasons is the evolutionary lack of facilities for authentic disguise as well as for the detection of lies. People in communities constantly lived close together and did not have many opportunities to cover cheating and the discovery of lies lead to the application of extreme sanctions. As we have seen in the study of Ekman (1996), none of these results showed the ability to accurately recognize lies. In this case, an appropriate method has proven to be the training by FACS

(Facial Action Coding System) allowing successful lie detection by examining emotional expressions in 70% of cases.

In [5] authors proposed a method for automatically detecting deceptive speech by relying on predicted scores derived from emotion dimensions such as arousal, valence, regulation, and emotion categories. They showed that, the emotional features have a relatively high predictive power in the deception task even when used by themselves. These findings imply that emotional attributes, even ones generated by machine learning systems trained on separate data, have considerable potential for detecting decep- tive speech. More generally, the results obtained support the feasibility of utilising labelled paralinguistic data in solving re- lated problems, for which annotated data is scarcely available.

Emotions play significant role in doctor's and patient's life. It can tell a much more information about health, different problems, psychological and physical condition of user interacting with our system.

1.5 Last Updates in the Lie Detection Methods

1.5.1 Latest Updates. In 2020, the ProfileCenter automated profiling module began to be used in the largest companies of the Russian Federation and even several English-speaking companies. Also, this module was the first to learn to determine the typical user behavior and its risks in the field of personnel, information and economic security.

Strengthening "neurologization" and the multimodal approach in lie detection.

Today, high-quality profiling and lie detection begins with an objective assessment of the state, properties and stigmas of the human nervous system. Neurointerfaces begin to enter densely into lie detection, making it better and more objective.

Commonly used methods in lie detection: polygraph, EEG, IT tracker, voice and emotional analyzers, speech and handwriting analysis, interviews, etc. Experts say that the multimodal trend will only intensify.

Many of the largest companies (including the Russian Federation) and some high-tech startups are accelerating efforts to create a multimodal remote lie detector.

1.5.2 Lie Detection and Artificial Intelligence. The system uses artificial intelligence - designed at Manchester Metropolitan University - to work out whether people are lying about who they are and why they are travelling. Dr Keeley Crockett from the School of Computing, Mathematics and Digital Technology, explains how it works. It's being showcased at the Science and Industry Museum as part of the Manchester Science Festival on Saturday October 20.

Their product is based on a non-verbal behavior with fine grained micro gestures, they don't consider such things as smiling or frowning, they are looking at a very small movements, such as an eyes moving left or right. They used these data and combine it together with the risk score and they use this in order to say for a particular question whether someone's lying or not. It was not just their product and their particular application that gives the wrist or it actually combined within the system to give an indication to the border guard. It is a human in the loop system. It doesn't make a full automatic decision it gives a risk score for each individual traveler.

2 RELATED WORK

Emotions are detected using different form of input, such as speech audio, text phrases or emoticons, facial expressions from video. These inputs can be derived from different form of applications. Recently, electroencephalogram (EEG) signal-basedemotion recognition systems are taken into account, though the use of EEG cap is invasive and may be uncomfortable to the users.

Thus, the most common input is combinations of speech and video. It can be collected in a non-invasive manner and more expressive than other forms. In [6] an audio-visual emotion recognition system using a deep network to extract features and another deep network to fuse the features. There are two modalities of input to the system: speech and video. Speech signals and video signals are processed separately and fused at the later stage before classification. There are two main steps for each of these modalities before fusion. The steps are preprocessing and deep networks using the CNN.

In [7] it is possible to extract basic algorithm of emotion recognition, even though it is using mobile application. In the paper it uses video from the phone, though we can use different forms of video capturing to get facial emotions. Then it is used to get most representing frames. There are many feature selection techniques, each of which has its own advantage and disadvantage. In their proposed system, authors adopt the KW technique for its simplicity and low computational complexity. The selected features are fed into a GMM based classifier. During training, models of different emotions are created from the feature set. The emotion corresponding to the maximum score is the output of the system.

In [8] propose a solution using texture descriptors, where descriptors use to be sets of numbers which indicate characteristics of the image. Local Binary Pattern (LBP) - a simple but effective texture descriptor which label every pixel of the image analysing its neighbourhood, studying if the grey level of every neighbour pixel is over a certain threshold and codifying this comparison by a binary number. This descriptor has become very popular due to its discriminative power and its low computational cost, which allows a real-time image processing. In addition, this descriptor has a great robustness in the present of grey level changes produced by light changes. After extraction features they use Support Vector Machine (SVM) to classify emotions.

In [9] authors described two methods of speech emotion recognition are described. The first one is about global statistics using Gaussian mixture models, and the second one is applying continuous hidden Markov models considering several states. The results of both engines reach the abilities of a human decider as described above. Results in recognition of seven discrete emotions exceeded 86% recognition rate.

This work [10] aims at showing improved performances of an emotion recognition system embedding information gathered from cardiorespiratory (CR) coupling.

For the [11] authors claimed that there has been a recent increase in the number of studies examining emotion recognition in different patients groups, however emotion recognition in patients with traumatic brain injury (TBI) is of particular interest because of the possible link with changes in emotional and social behaviour.

Changes in emotional and social behaviour are relatively common following TBI.

In [12] research authors state that the monitoring of elderly people on an individual basis, in a medical sense, will not be a viable proposition in the future. The infrastructure available is not adequate to meet all expectations and subsequently people will continue to live at home with inadequate care. In the Future, more experiments will take place in order to enhance the emotional model and to add more emotions such as sadness, and fear to the agent.

The similar ideas are presented in [13] developes a system relied upon a novel emotion modelling methodology comprising of a number of research steps aiming to address the challenges associated with a medical environment, and the complex nature of patient's emotions. This emotion modelling methodology provides a number of ideas contributing to the development of effective healthcare computing applications.

The main goal of [14] is to gradually replace the use of complex techniques, such as Electroencephalogram (EEG), and to introduce a new way of approaching rare disease using a simple desktop or laptop computer and a webcam. They provide medical researchers with a new, non-invasive and portable analysis system to study progressive cognitive degradation through the interpretation of shown facial emotions.

3 USER STUDY

For the user studies we were interviewing several doctors from Innopolis and Kazan cities. In Kazan we have interviewed allergist, therapist and gynecologist, in Innopolis - sports medicine doctor. We were interviewing them personally, asking about their general work and usual working day, which problems and inconvenience they face every day. The biggest problem for them during their job in the government clinics were lack of electronic patient management system, because of that they spend a lot of time by handwriting. In private clinics there are comfortable electronic systems, however it would be more convenient to have a speech recognition system and, accordingly, its recordings. They also put forward the idea of a system that would help nurses transfer or move patients who are unable to move independently. Especially heavy patients.

Speaking of gynecology, the idea of creating a transformable bed for women in labor was discussed, since each of them has different parameters, gynecologists and their assistants manually twist the beds during childbirth, and sometimes there's absolutely no time for this.

Also some kind of the "smart home" system with adjustment of light, temperature, humidity and so on, in hospital wards could be helpful.

They also mentioned that, as a rule, the greatest difficulties are caused by the emotional state of doctors, especially in severe cases.

One of the doctors from Kazan speaking about our idea remembered about "da Vinci Surgical System", considering something similar as a possible future extension of the system.

Sports medicine doctor from Innopolis commented told us that emotion recognition system could be very helpful for leaving only valid symptoms and that it will also help to assess the degree of pain and its effect on the patient's standard of living.

Moreover we talk with student of the last year of the medical University in Ufa city to get her opinion about Intelligent framework for utilizing facial emotions. She considers such system interesting and useful in different medical specializations.

4 RESULTS

Medicine is a special kind of environment where there are a lot of communication between people. Doctors talks with patients to retrieve as many as possible information about their health to treat them right. Doctors should extract only valid symptoms from patients. It can be done via analysing emotions during treatment session. Medicine is similar with others fields like the field of criminology where facial expressions are actively used.

We analysed different prototypes to choose the most convenient one for medical personal. The final version is a web application with only two screens. The first one is to control recording, adding some metadata about patients, managing questions and aligning the answers with respect to question time periods. The second one is for deep understanding emotional model of patient during the answering on particular question. Doctor can select any question from list of questions with answers on the right side of screen. The system can automatically detect the pattern of changing emotions in a way corresponding to lie if it occurs and then notifies the doctor. Also, it allows assessing the degree of pain and its impact on the patient's standard of living.

As a results of the user study (interviews with different types of doctors) as part of our research their opinion was that this system can be a good help for a doctor, especially with problematic patients or patients who come for consultations without signs of illness, thinking up illnesses for themselves.Also, such a system will help to leave only valid symptoms. It will help to assess the degree of pain and its impact on the patient's standard of living. A common problem in the practice of doctors is when they point to a specific point in the patient's body with the question of whether he is experiencing pain in this place, and if he is experiencing which kind of pain is this and how strong the patient cannot understand and give an assessment. Therefore, in this situation, a system that helps and at least somehow evaluates emotions would be very useful. Moreover, such a system would be useful, because doctors do not have good or positive questions, all questions in one way or another cause stress or discomfort for the patients. Therefore such situations can cause mild panic in patients, which in turn leads to confused and untrue responses.

Since the recording of patient answers on the video is private process, patient will be provided several ways of consultation: recording the consultation with the user permission to store consultation process and results, recording the consultation without storing the recorded data, and the last option is consultation without any recordings. It is important for doctor to present the aim of recording correctly, since confusion may arise on the patient side. Doctor should neatly explain the pros of options which are include the recording process: the first is that storing the data will provide the ability to process it and, therefore, improve the accuracy of the system, and the second is that using the system helps to doctor to better understand patient, hence, to make a more accurate diagnoses.

5 CONCLUSION & FUTURE WORK

We propose a project called "EMO" which represents system using deep neural networks for utilizing facial emotions of patients from videos produced by usual cameras like phone, tablet or surveillance cameras. It can evaluate how positive and brightly expressed emotions of patients are in different scenarios.

For the future work such system can be trained on others emotion patterns to automatically detect not only lie but the patterns corresponding to different medical diagnoses. Also, the system can be used to assess feelings of medical personal to interrupt them from different working activities if they are tired or emotionally unstable.

REFERENCES

[1] Paul Eckman. Universal and cultural differences in facial expression of emotion. In *Nebraska symposium on motivation*, volume 19, pages 207–284, 1972.

[2] Why people lie?

[3] Lies in the doctor-patient relationship. 2009.

[4] Martina Vicianova. Historical techniques of lie detection. *Europe's journal of psychology*, 11(3):522, 2015.

[5] Shahin Amiriparian, Jouni Pohjalainen, Erik Marchi, Sergey Pugachevskiy, and Björn W Schuller. Is deception emotional? an emotion-driven predictive approach. In *INTERSPEECH*, pages 2011–2015, 2016.

[6] M Shamim Hossain and Ghulam Muhammad. Emotion recognition using deep learning approach from audio–visual emotional big data. *Information Fusion*, 49:69–78, 2019.

[7] M Shamim Hossain and Ghulam Muhammad. An emotion recognition system for mobile applications. *IEEE Access*, 5:2281–2287, 2017.

[8] Francisco A Pujol, Higinio Mora, and Ana Martínez. Emotion recognition to improve e-healthcare systems in smart cities. In *The International Research & Innovation Forum*, pages 245–254. Springer, 2019.

[9] Björn Schuller, Gerhard Rigoll, and Manfred K. Lang. Hidden markov model-based speech emotion recognition. In *ICME*, pages 401–404. IEEE Computer Society, 2003.

[10] Gaetano Valenza, Antonio Lanata, and Enzo Pasquale Scilingo. Improving emotion recognition systems by embedding cardiorespiratory coupling. *Physiological measurement*, 34(4):449, 2013.

[11] Magdalena Ietswaart, Maarten Milders, John R Crawford, David Currie, and Clare L Scott. Longitudinal aspects of emotion recognition in patients with traumatic brain injury. *Neuropsychologia*, 46(1):148–159, 2008.

[12] Sherief Mowafey and Steve Gardner. A novel adaptive approach for home care ambient intelligent environments with an emotion-aware system. In *Proceedings of 2012 UKACC International Conference on Control*, pages 771–777. IEEE, 2012.

[13] Faiyaz Doctor, Charalampos Karyotis, Rahat Iqbal, and Anne James. An intelligent framework for emotion aware e-healthcare support systems. pages 1–8, 12 2016.

[14] V. Bevilacqua, G. Grimaldi, D. Semeraro, S. Simeone, M. Delussi, and M. de Tommaso. An innovative framework for rare neurodegenerative diseases analysis. In *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, pages 555–559, May 2015.