Research Methods draft

Mario Tsatsev s1028415

May 2020

1 Introduction

Stress management is related to public healthcare and quality of life, an accurate stress classification method is necessary for the design of stress monitoring systems. Why ? Current methods like measuring heart rate or regular visit to psychologists are costly time-wise and impractical for most of the common professions out there. Although fire-fighters, policemen and military personnel are the priority targets of stress monitoring, a paper[1] from the "European Agency for Safety and Health at Work", about the costs of work related stress and its treatment reveals tremendous expenses for the mentioned in the read European countries, which shows that this is a problem in the overlooked office jobs as well. Such systems could be a machine learning and more specifically deep learning algorithm[2][3][4][5] which can be build into the computers of the company. This method can be applied continuously throughout the day, constantly monitoring the emotion behavior of the subject. In addition, traditional methods lack the ability to measure stress over long periods of time without disturbing the worker's routine but possibly with this software the person in front of the computer will barely even notice that it is there.

2 Problem formulation

For a long time now a "war" on stress has been waged by psychologists and others. One of the main "battlefields" is the causes of stress. Typically we all agree that the work place is an area that is a common cause of stress. Identifying physical symptoms of a range of conditions is something a psychologist does, but if he can, can't machines do it as well ? In fact can they do it better ? Can they monitor healthcare parameters in a scalable low cost ways. If it is continuous then it is scalable, while a specialised persona might be limited by time but also range of observations. It will be hard to look at the mental state of many individuals all the time, at the same time.

... work place is an area that is a common cause of stress ...

A website with a trustworthy name agrees with me [8]

3 Research Question

Can we use an emotion recognition deep learning system to monitor employees stress at work ?
But also:

3.1 Questions related to the overall system

- How to built it as an embedded software ?
- Which types of neural networks are useful ?
- Which type of learning will be applied ?
- How to gather enough data ?
- Data acquisition from speech ?
- Data acquisition from video frames of the face and upper body ?
- How to specialize the network ?
- Why would this be better than what we already have ?
- What are the expected final results ?

4 Motivation

First and foremost I am really interested in machine learning and would like to test its limits and implementation as much as possible. Researchers are always looking to do new things by building on what has been done before. This is not the first time such a system is proposed but as far as I know, it is the first time when it will be applied in such a way. Interestingly this is not a new thing. Methods to monitor stress, discouragement and inattention have been around since 2003[7], however with the resources they had, their scope was limited to more "serious" jobs, like personal in nuclear power plants.

5 Data generation methods

This section is partially related to one of the questions, namely "How to gather enough data ?". It can be and in fact must be conducted before and after the implementation of the algorithm. Four steps have to be taken:

- Choose the respondents: The research is about using machine learning to monitor stress, however for that we need to know what stress is. For that we need to contact specialized person in that field, a psychologist. It would be helpful to get the expertise of more than one psychologist.
- Setting the interview: This one is hard to think of as it may vary for every respondent. Ideally in person interview is preferred over handing questionnaires, since the questions will likely change based on the answers for each of those, as well as further clarification could be requested if it is felt necessary.
- Kind of questions: Assuming we can build the model that can recognize different facial expressions and clearly makes a difference between various tones of the complexion, tone of voice, amount of moves made on the video per frame, isolate itself only to real human voice and not get distracted by music etc. We need to tell the algorithm what to infer from those classifications. So this is what our questions have to be about.
- Example questions:
  - How does our voice change when we are under stress ?
  - How does the speed of our speech change when we are under stress ?
  - Are there patterns of words present more often when we are under stress ?
  - Are there any behavior-related patterns that tell us that an individual is under stress, maybe scratching his head ?
  - When does the change in our voice tell us that the individual had too much ?
  - Can our complexion tell us something about our mental well-being and what are the signs to look for if we want to measure stress level ?
  - How accurately do facial expressions measure stress level ?
  - As a psychologist, do you feel fine working with your patient's digitally generated stress profile ?
  - If you are fine with that, what do you expect to see on such a psychological profile ?

6 Theoretical framework

In order to answer the research question(s), we need a theoretical framework to build our research upon. One of the building blocks of the theoretical research will be the reasoning about the implementation of the algorithm. As discussed before such algorithms are not new and researchers like to build new things by improving on others which already exist. In the reference page there are four

AI algorithms[2][3][4][5] which attempted to build such a model using either video which includes eye tracking or facial expressions but not both, audio or both. These papers will be further reviewed in more detail and will likely become our stepping stone and baseline. Meaning that before we put this algorithm in practice is should be at least sufficient to get equal and hopefully better results on the training and testing data sets as the upper mentioned work. This can easily be achieved since their data sets are publicly available! Building a framework for the psychological analysis of what the network would infer from the subjects which it will be tested on is at this moment, impossible for me. I lack specialised psychological knowledge.

7 Timeline

This research might take too long. We need to divide the available time into research phases. After the phases have been explained a timeline will be made to divide the available time between the various phases.

7.1 Training data

TBD

7.2 Algorithm

TBD

7.3 Interview

TBD

7.4 Test of the algorithm in real environment

TBD

8 Literature

Here I have mainly listed in the reference section, works related to the computer science part of the research. This includes various attempts at building an optimal model for emotion recognition, gathering data or implications of "emotion aware computing". This also means that I have not read or looked at any psychology papers or publications on stress or in general.

9 Reference

[1] Calculating the costs of work-related stress and psychosocial risks – A literature review
[2] Hossain, M. Shamim Muhammad, Ghulam. (2018). Emotion Recognition Using Deep Learning Approach from Audio-Visual Emotional Big Data. Information Fusion. 49. 10.1016/j.inffus.2018.09.008.
[3] C. Aracena, S. Basterrech, V. Snáel and J. Velásquez, "Neural Networks for Emotion Recognition Based on Eye Tracking Data," 2015 IEEE International Conference on Systems, Man, and Cybernetics, Kowloon, 2015, pp. 2632-2637, doi: 10.1109/SMC.2015.460.
[4] Deep Learning based Emotion Recognition System Using Speech Features and Transcriptions
[5] ResNet-50 and VGG-16for recognizing Facial Emotions
[6] AINow 2019 Report
[7] M. Pantic and L. J. M. Rothkrantz, "Toward an affect-sensitive multimodal human-computer interaction," in Proceedings of the IEEE, vol. 91, no. 9, pp. 1370-1390, Sept. 2003, doi: 10.1109/JPROC.2003.817122.