SENIOR DESIGN PROJECT

Horus

Project Specifications Report

Ufuk Palpas, Furkan Demir, Can Kılıç, Asya Doğa Özer, Muzaffer Köksal

Supervisor: Uğur Güdükbay
Jury Members: Shervin Arashloo, Hamdi Dibeklioglu

October 11, 2021

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project cours CS491/2

1.Introduction

In the 21st century, human-machine interactions were researched through speech. Unfortunately, these researches led to a dead-end "because the machine does not understand the emotional state of the speaker. Thus, researchers have introduced a newer research field compared to others, emotion recognition" [1]. Emotion recognition is the process of identifying human emotion [2]. People can have various amounts of success in identifying human emotions. For computers, emotion recognition involves face detection that leads us to the computer vision area. Computer Vision is one of the most popular computer science topics of the last decade. New technologies of computer vision such as face recognition, defect detection, intruder detection, etc. are highly invested in these days. There are a lot of applications of emotion recognition. However, there are no significant examples of multi-person emotion recognition. Our project will focus on that area, combining emotion recognition with human face recognition in crowded areas.

Furthermore, Horus will also have a deception detection mod. To obtain such functionality, we will use emotional cues described in psychology articles. Some research shows that emotional cues can be used as a diagnostic tool to detect deception [3]. We will use these emotional cues decoded through Horus to detect deception.

This report will contain information about the Horus, its constraints, professional and ethical issues, functional and non-functional requirements.

1.1 Description

The goal of Horus is to recognize human emotions from the mimics of their faces. A camera will rapidly take images/video and Horus will analyze their faces. In single user mode for Horus to analyze a user's face, there shouldn't be any obstacle, so if the user wears a mask, hat, etc. Horus will warn the user and will be able to analyze the emotions of the user in real time besides emotion analysis of a saved video or image. It will report the various emotions detected from the subject’s face and the intensity of said emotions on a scale. It will also attempt to detect the mood of the subject from long term analysis of his emotions.

Horus will also have a crowd control mod where it will analyze the emotions of a group of people such as stand-up, theater or movie audiences to provide feedback to show organizers who can look at audience excitement and decide which content to keep and which to discard. A playwright, for example, can look at the reception to specific acts of his play and decide on which changes to make. As another example, a street artist can place a camera nearby and get feedback from the Horus according to the emotions of the audience.

Another use case can be handling tension and preventing aggression. For example, Horus can be used in prison common areas to keep an eye on prisoners. If Horus detects multiple nervous, angry prisoners and it alerts the guards.

1.2 Constraints

1.2.1 Implementation Constraints

- The Github platform will be used to orient and control the progress of the project.
- Horus will be a desktop platform application that can be connected to the smartphone for giving feedback and notifications.
- OpenCV [4] face detection library and FER-2013 [5], CK+ [6], FERG [7] and JAFFE [8] datasets will be used for emotion detection.
- Fer (facial expression recognition) python library will be used [9].
- CrowdHuman [10], WWW Crowd [11] datasets will be used for human face detection in multi-person scenes for crowd control mod.

1.2.2 Economy Constraints

- Many of the libraries and frameworks are free to be used for research.
- FER-2013, CK+, FERG and JAFFE datasets are free to be used for research.
- Google Colab is a free service that Google offers which provides a 12 hour continuous GPU and TPU if our GPUs are not enough for the training process.

1.2.3 Sustainability Constraints

- User experience will be taken into consideration for updating Horus.
- The results of the face analysis will depend on the initially provided data. It will continue analyzing until the provided data ends.

1.2.4 Language Constraints

- Since Horus detects emotional cues, language does not matter.
- The audible warnings will be in English.

1.2.5 Ethical Constraints

- Personal information will not be shared via any third-parties.
- Recorded videos or images of crowded areas will not be published or stored not to violate personal rights and ethical issues.
- Our application will adhere to ethics by the National Society of Professional Engineers.

1.2.6 Security Constraints

- Collected data will be stored in the user's own storage area and Horus will not use an internet connection, thus the data will be secured.

1.3 Professional and Ethical Issues

Undoubtedly, as an application that uses visual data and being a computer vision project there are various professional and ethical issues. These issues are prominent for both the development and usage of the application.

We will consider several professional and ethical issues during and after the development process of Horus. One of them is data security. We have to get video input from the user to do emotion recognition and that video input should not be shared with any kind of third party. Furthermore, we will ensure that the stored video and image data should not be leaked at all. We will also ensure that Horus will not be used to spoil and put pressure on any user unless there is a security concern.

For the sake of professionality, the group meetings will be arranged each week. What has already been done and what will be done until the next meeting will be determined and implemented in the closest time span. The source code of the horus will be private and stored on Github servers. Contribution of the group members will be one of our main concerns.

2. Requirements

2.1 Functional Requirements

2.1.1 System Functionalities

- The system should give an opportunity to the user to select what will be analyzed, such as image, video or real time analysis.
- The system should request permission to access the local storage of the user to access videos or images.
- The system should request permission to access the camera of the user’s computer.
- The system should give feedback to the user according to the analysed data which contains the emotions of the surrounding people, with the aid of an audible warning.

2.1.2 User Functionalities

- The user can choose between single user mode and crowd control mode.
- In the single user mode, the user should see the results on the same page with the analyzed data.
- In the crowd control mode, the user will get feedback from Horus according to the crowd's emotions.
- The user can connect to Horus via smartphone and earphones to get feedback remotely in crowd control mod.

2.2 Non-Functional Requirements

2.2.1 Reliability

- Horus will detect the emotions as correctly as possible.
- Horus will determine the deception as much as possible.

2.2.2 User-Friendliness

- The user interface of our app must be simple and easily understandable.
- It shouldn’t require any pre-education to use it.
- User can get the audible feedback easily from his/her smartphone.

2.2.3 Efficiency

- Horus will use the resources (such as CPU, RAM etc.) of the computer which it runs on to give accurate results in real time without any delay to give immediate feedback to the user.
- It should detect the general mood of the user with a long enough video.
- In crowd control mod, Horus will calculate the average emotion of the crowd.

2.2.4 Security

- The photos/videos that our app captured will not be shared via any third parties by Horus.

3. References

[1] J. Nicholson, K. Takahashi, R. Nakatsu, “Emotion recognition in speech using neural networks" *Neural Comput. Appl.*, 9 (2000), pp. 290-296

[2] “Emotion Recognition.” *Wikipedia*, Wikimedia Foundation, 21 June 2021, en.wikipedia.org/wiki/Emotion_recognition. [Accessed 09/10/2021].

[3] Zloteanu, Mircea. “Emotion Recognition and Deception Detection.” *PsyArXiv*, 25 Sept. 2019. https://psyarxiv.com/crzne/.[Accessed 09/10/2021].

[4] “Home.” *OpenCV*, 16 June 2021, opencv.org/ [Accessed 10/10/2021].

[5] Sambare, Manas. “Fer-2013.” *Kaggle*, 19 July 2020, www.kaggle.com/msambare/fer2013 [Accessed 10/10/2021].

[6] 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," *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.

[7] *Ferg-DB*, grail.cs.washington.edu/projects/deepexpr/ferg-2d-db.html [Accessed 10/10/2021].

[8] Lyons, Michael, et al. “The Japanese Female Facial Expression (Jaffe) Dataset.” *Zenodo*, 14 Apr. 1998, zenodo.org/record/3451524#.YWM4x9pByUk. [Accessed 10/10/2021].

[9] “Fer.” *PyPI*, https://pypi.org/project/fer/ [Accessed 10/10/2021]

[10] “CrowdHuman” https://www.crowdhuman.org [Accessed 10/10/2021]

[11] https://amandajshao.github.io/projects/WWWCrowdDataset.html