Bachelor Thesis Proposal

Comparative Analysis of Action Unit Recognition Methods for Autism Spectrum Condition Detection

Sebastian Jack Maier

December 2023

Duration: January - April 2024
1st Supervisor: Prof. Hanna Drimalla, Bielefeld University
2nd Supervisor: William Saakyan, Bielefeld University

1 Introduction

Facial expressions play a fundamental role in human communication, serving as a natural mechanism through which individuals discern and interpret specific emotions displayed on a person's face. This skill is a crucial aspect for communication, empathy, and trust between individuals [1, 2, 3]. Facial expressions can differ, both in variety and intensity, among individuals with distinct psychological conditions, like depression, Autism Spectrum Condition (ASC), and social anxiety. My thesis will focus on the *automatic detection of facial expressions* associated with Autism Spectrum Condition (ASC) - a condition characterized by challenges in accurately interpreting and responding to facial expressions [4, 5].

There are research efforts to create tools for quantitatively measuring facial expressions. Specifically, Action Units (AU) have been used to describe muscle movements in the lower and upper half of the face [6, 7]. With a combination of such AUs, emotion expression can be described as changes in facial muscle movements. Such quantitative descriptions can then help to distinguish conditions, in which the emotion expression might be different. For example, when an individual suddenly shows sadness, we would expect the conversation partner to react with a change in their behaviour and facial expression. An individual with ASC, however, might react delayed or not at all [5].

ASC is often difficult to diagnose, especially in adult individuals. Many grown ups have learned to cope and mask their lack of social skills. This makes it challenging to diagnose correctly and objectively. Additionally, most studies evolve around children with ASC, and have a significant bias towards male subjects [8]. To address this issue and to help the long diagnosis process, machine learning algorithms could be used in addition to medical staff [5].

This thesis builds upon previous efforts toward ASC detection using machine learning methods in combination with a standardized data collection protocol. Specifically, the dataset was collected using the Simulated Interaction Task (SIT), a standardized test in which participants communicate with a prerecorded interaction partner [9]. With recordings of study participants (ASC or neurotypical), it is possible to extract multiple non-verbal clues [10]. This includes gaze behavior, head pose, voice features, as well as facial expressions which we will focus on with Action Units.

The accuracy of detecting ASC with Facial Expression on the collected dataset was 73% with OpenFace 2.0 (2018) [10]. This performance could potentially improve by applying newer methods that surpass OpenFace in accuracy. OpenFace [11] is widely used as an open-source tool to detect Action Units since it came out in 2018. However, other tools and approaches of detecting Action-Units have emerged in recent years [7].

2 Task / Goal

The objective of this study is to assess and compare the performance of newer open-source models, released since 2020, in detecting Action Units, in comparison to the widely used OpenFace model [11]. Later the methods will be compared in the task of classifying Autism Spectrum Condition based on facial expressions.
Following criteria for the models must apply:

- Up to date:
  The methods should be open-source and from 2020 or newer, which means they should be newer than OpenFace[11].

- Performant:
  The method evaluation from the literature should demonstrate higher performance than OpenFace on at least one Dataset. No direct comparison with OpenFace is required, as long as the reported accuracy (or other metric) is better than that of OpenFace.

3 Literature & Methodology

Relevant Literature will be found by searching for specific keywords as 'Autism Spectrum Condition' and 'Action Units', while applying a date filter 'since 2020'. Following tools should help to find literature and methods: *Google Scholar* (scholar.google.com), *paperswithcode.com* and other specialized websites dedicated to scientific papers that may include open-source materials. Additionally, relevant references and comparisons in the identified works will be further examined.

The relevant methods should claim to perform better in detecting AUs than OpenFace for at least 50% of AUs, as different Action Units have different base accuracy [12]. Additionally, the average performance, based on Pearson Correlation Coefficient (PCC) between the predicted Action Units and the ground truth or similar metrics, should be higher than OpenFace. If no comparison with OpenFace could be found in a specific paper, the overall performance should still be better than the average performance of OpenFace.

In total, three models should be compared towards OpenFace.

4 Evaluation

4.1 Feature Extraction

Suitable models will be applied to the SIT dataset. Initially, these models will be tested on a test video to evaluate the usefulness of the features for the specified task and to assess their comparability with OpenFace features. The steps taken to execute it will be documented along with any modifications made. After that the actual feature extraction will be run on the intern server to maintain data anonymity. The result will be saved for further evaluation and the classification task.

4.2 Feature Analysis

The next step will be to analyse the extracted features in terms of both intensity and occurrence. The intensity of AU's will be plotted and compared for every method to see the difference in detection and it will be checked if the values correlate between those methods. Similarly, the occurrence of the AUs will be compared using binary values.

4.3 Classification Task

My hypothesis is, that I can detect improvement on accuracy, recall and precision regarding the task of detecting ASC. The extracted Action Units or features will be used to train XGBoost [13] in a similar manner as done by Saakyan et al. [10]. The result will be a classification prediction of given participants, labeled as ASC or NT. A table will be created to compare every applied method and OpenFace data using cross-validation on the classification task. Additionally, I will visually present the results, including plotting ROC curves of each method to illustrate their performance.

4.4 Action Unit Detection

Additionally, I will trace back and analyse the decision process further. Important AUs for detecting ASC will be selected based on related literature. In particular, Action Units related to positive emotions as AU6 and AU12 tend to be less present on individuals with ASC compared to neurotypical persons [14],

so they should be examined.
SIT contains three different parts, including the parts, where particpant is expected to have positive emotions and the part with expected negative emotion - digust. Time stamps are provided to enable grouping of expected expression. I will use this to create figures to visualize detection of certain AUs in different parts and compare output of the methods.

Visualization should include diagrams for each relevant AUs, where the intensity detected by every method will be made visible over time. With this, the sensitivity of a method regarding the respective AU can be identified. Furthermore, the most important set of AU's for the differentiation between neurotypicals and individuals with ASC will be analysed using SHAP (or similar) library to find most relevant features and the classification performance will be compared between different models.

The underlying code for feature extraction and analysis will be provided in a frozen GitHub repository. The code should be sufficiently documented and match the results described in the bachelor thesis. The efficiency of the code will not be part of the evaluation.

4.5 Risks and Uncertainties

Given the uncertainties associated with an available open-source methods, several potential risks have been identified and will be addressed in the following. The first potential problem might be reported metrics. Some methods report accuracy metric for AU detection and others Pearson correlation coefficient. If the method can not be compared with OpenFace directly or through other methods, which report both metrics, the method will be still included if it is newer than OpenFace. The other potential problem might be limited open-source method availability or poorly written code. If code is available but it can not be executed after 20 working hours, the method will be excluded. In case there are less than three open-source methods available and suitable for the task or any other problem will occur during the evaluation and extraction part, next steps with comparable load should be additionally done:

- Other descriptors/functionals will be used to analyse and classify the data than what was done in [10].
- I will record a video of myself displaying various facial expressions, comparing the expected outcomes with the extracted features by plotting the intensities of Action Units. This process will help in assessing the sensitivity and performance of the methods in a video captured under conditions similar to those of the SIT studies.

References

[1] Hillary Anger Elfenbein and Nalini Ambady. On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128(2):203–235, 2002. Place: US Publisher: American Psychological Association.

[2] Jean Decety and Philip L. Jackson. A Social-Neuroscience Perspective on Empathy. *Current Directions in Psychological Science*, 15(2):54–58, April 2006. Publisher: SAGE Publications Inc.

[3] Stephen Porter, Leanne ten Brinke, and Brendan Wallace. Secrets and Lies: Involuntary Leakage in Deceptive Facial Expressions as a Function of Emotional Intensity. *Journal of Nonverbal Behavior*, 36(1):23–37, March 2012.

[4] DSM. *Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed.* Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed. American Psychiatric Publishing, Inc., Arlington, VA, US, 2013. Pages: xliv, 947.

[5] Kaushik Vakadkar, Diya Purkayastha, and Deepa Krishnan. Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques. *SN Computer Science*, 2(5):386, July 2021.

[6] Ruicong Zhi, Mengyi Liu, and Dezheng Zhang. A comprehensive survey on automatic facial action unit analysis. *The Visual Computer*, 36(5):1067–1093, May 2020.

[7] Felipe Zago Canal, Tobias Rossi Müller, Jhennifer Cristine Matias, Gustavo Gino Scotton, Antonio Reis de Sa Junior, Eliane Pozzebon, and Antonio Carlos Sobieranski. A survey on facial emotion recognition techniques: A state-of-the-art literature review. *Information Sciences*, 582:593–617, 2022.

[8] Luca Zaniboni and Jørn Toftum. Indoor environment perception of people with autism spectrum condition: A scoping review. *Building and Environment*, 243:110545, September 2023.

[9] Hanna Drimalla, Tobias Scheffer, Niels Landwehr, Irina Baskow, Stefan Roepke, Behnoush Behnia, and Isabel Dziobek. Towards the automatic detection of social biomarkers in autism spectrum disorder: introducing the simulated interaction task (SIT). *npj Digital Medicine*, 3(1):1–10, February 2020. Number: 1 Publisher: Nature Publishing Group.

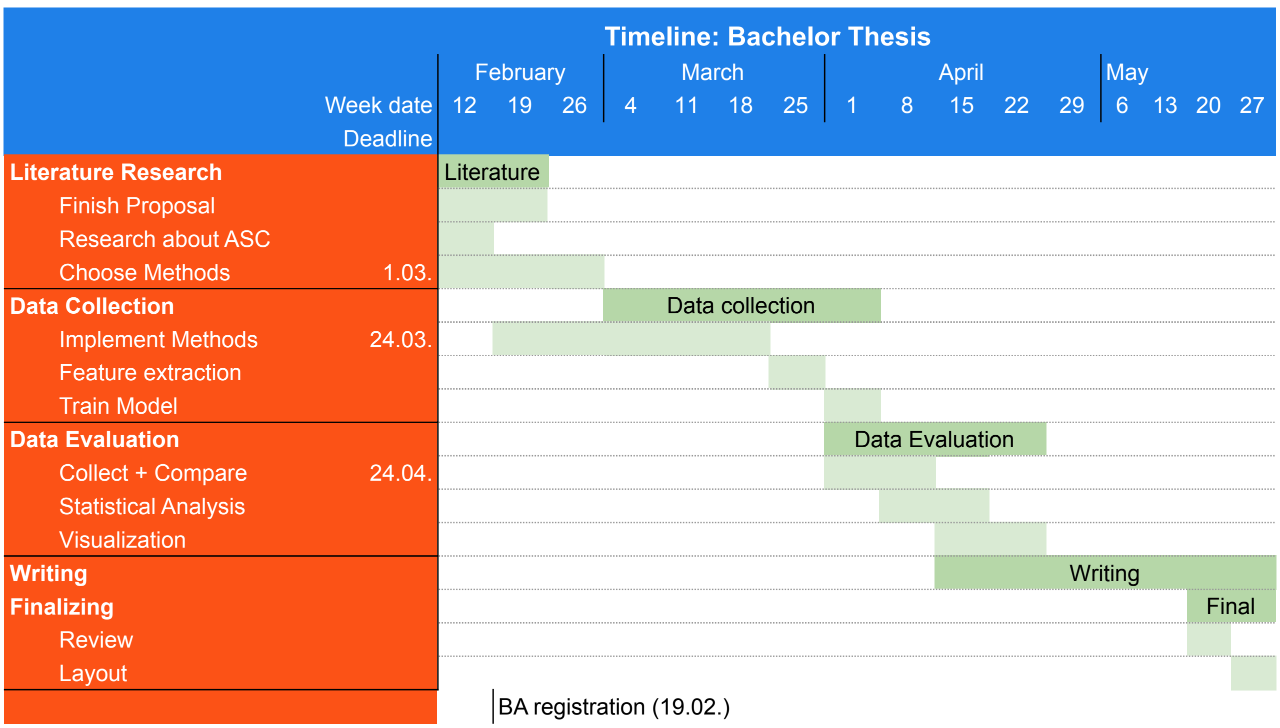
[10] William Saakyan, Matthias Norden, Lola Herrmann, Simon Kirsch, Muyu Lin, Simon Guendelman, Isabel Dziobek, and Hanna Drimalla. On Scalable and Interpretable Autism Detection from Social Interaction Behavior. *11th International Conference on Affective Computing and Intelligent Interaction*, 2023.

[11] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. OpenFace 2.0: Facial Behavior Analysis Toolkit. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 59–66, May 2018.

[12] Di Chang, Yufeng Yin, Zongjian Li, Minh Tran, and Mohammad Soleymani. LibreFace: An Open-Source Toolkit for Deep Facial Expression Analysis, August 2023. arXiv:2308.10713 [cs].

[13] XGBoost Documentation — xgboost 2.0.2 documentation.

[14] Abigail Bangerter, Meenakshi Chatterjee, Joseph Manfredonia, Nikolay V. Manyakov, Seth Ness, Matthew A. Boice, Andrew Skalkin, Matthew S. Goodwin, Geraldine Dawson, Robert Hendren, Bennett Leventhal, Frederick Shic, and Gahan Pandina. Automated recognition of spontaneous facial expression in individuals with autism spectrum disorder: parsing response variability. *Molecular Autism*, 11(1):31, May 2020.


Timeline: Bachelor Thesis
February
March
April
May
Week date
12 19 26 4 11 18 25 1 8 15 22 29 6 13 20 27
Deadline
Literature Research
Finish Proposal
Research about ASC
Choose Methods
1.03.
Data Collection
Implement Methods
24.03.
Feature extraction
Train Model
Data Evaluation
Collect + Compare
24.04.
Statistical Analysis
Visualization
Writing
Finalizing
Review
Layout
Literature
Data collection
Data Evaluation
Writing
Final
BA registration (19.02.)