Intelligent Interactive Systems (1MD039)
Uppsala University – Autumn 2024
Assignment 1 reflections by Group 1

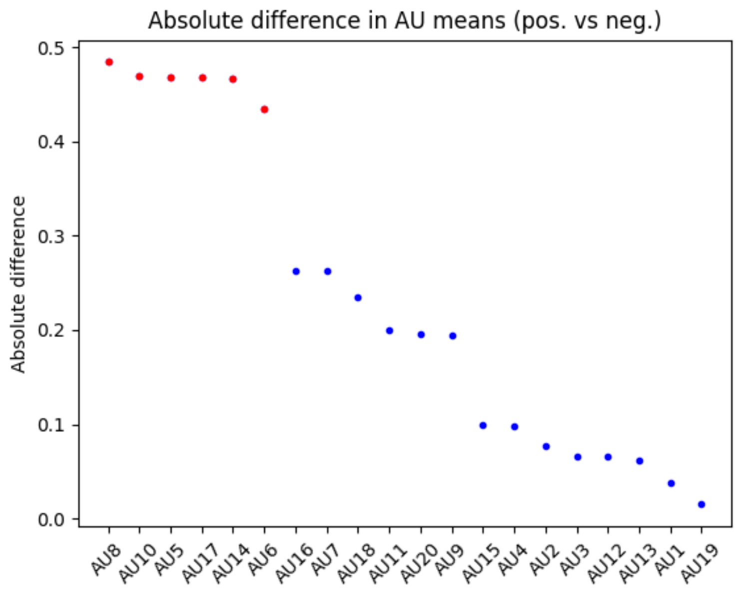
Nima Dehghanian and Axel Gleisner

22nd November 2024

1 Py-Feat detector

1.1 Predictions based on visualization

We agree with most of the predictions made by the Py-Feat detector. The system generally performs well in identifying facial features and extracting AU activations and expressed emotions. However, it does appear to encounter issues when dealing with occluded faces or certain angles. This is to be expected, and we believe Py-Feat to be well within the threshold of acceptable limits.



1.2 Points of failure in detection

If a face is at an extreme angle or obscured in some way, as seen in the first two images in A.1, there will often be difficulties with detection. This is to be expected. The detector scans for positions of facial features and is able to handle when some are left out, if the rest of the face receives a high enough probability score. There is no scan being performed for bodies, and thus the position of a face cannot be extrapolated from the position of any body seen in the image, this is evident in the second image of A.1. As for incorrectly identifying faces where there are none - this is rare but can occur in some circumstances, as in the last image of aforementioned appendix images. Other points of failure can occur regarding expression detection. These misinterpretations are often due to a blurred face's expression, or an angled face (see A.2).

1.3 Py-Feat compared to human observer

In the provided dataset there seem to be no major obstacles for a human observer sans cases where the face is completely obscured or angled away. In software that solely identifies faces by facial features it is to be expected that a human perspective would be superior. In order to get closer to human precision the software would need to take context clues into consideration, such as scanning for a persons body or hair to extrapolate where a face would be.

2 AU activations

2.1 Inputs for predictive algorithm

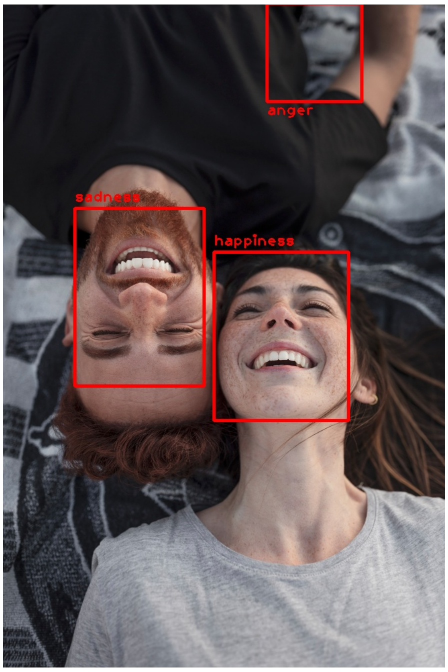
For the analysis we have performed, we would choose the red subset of AUs in figure 1.1 as inputs for a predictive algorithm. AUs with high absolute mean differences are good candidates to distinguish between different expressions as they provide the most significant variation in emotional state.

2.2 Overfitting

The problem with using too many features is that it can lead to overfitting, where the model learns noise in the training data rather than the underlying patterns. Overfitting reduces the model's ability to accurately assess new, unseen data. Additionally, using too many features can increase the computational complexity and make the model more difficult to interpret. Therefore, selecting a subset of the most relevant AUs can improve the model's performance and ease of working with said model.

A Appendix

A.1 Face detection failure



A.2 Expression detection failure