Assignment 2

5XSL0: FUNDAMENTALS OF MACHINE LEARNING

Dr. Ir. Luis A. Zavala Mondragón

Prof. Fons van der Sommen

Prof. Uzay Kaymak

Cover image generated with MS365 copilot ver. 19.2509.54051.0

Dept EE – Signal Processing Systems group

Version Sept 2025

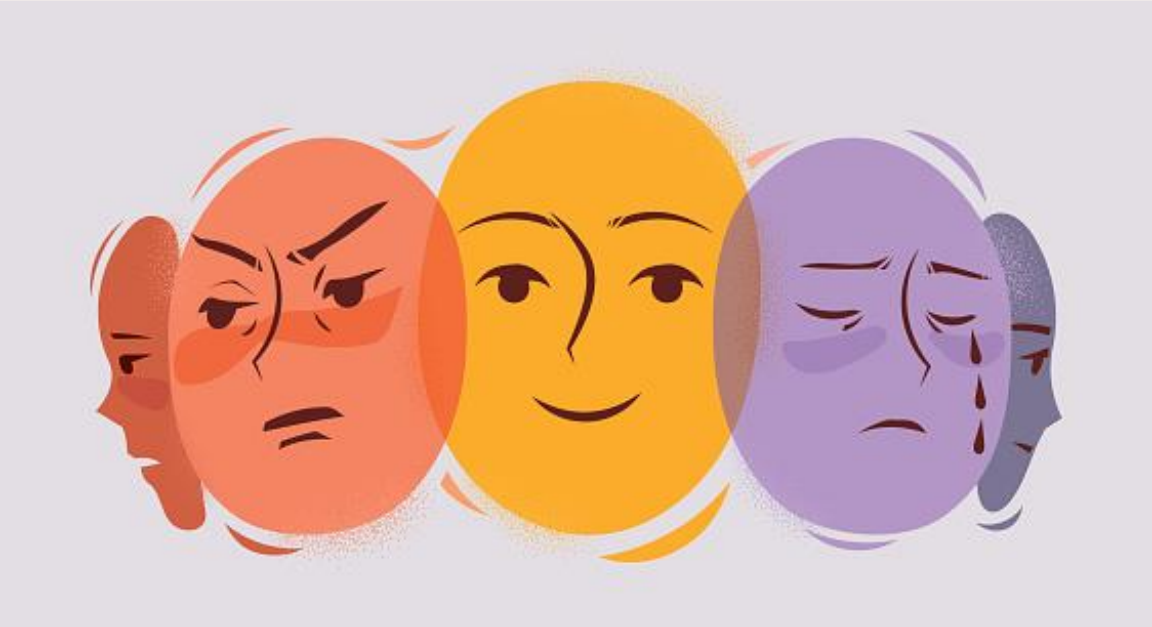
Introduction

- **In the course we have seen different methods for building machine learning models.**

- **For assignment 1 you built a pipeline in which you had to acquire data, perform feature extraction and also develop models.**

- **This assignment will test other types of signals and you will face other issues when building machine learning models**

This assignment …



Source: Various emotions and facial expressions of one person. by Woocat

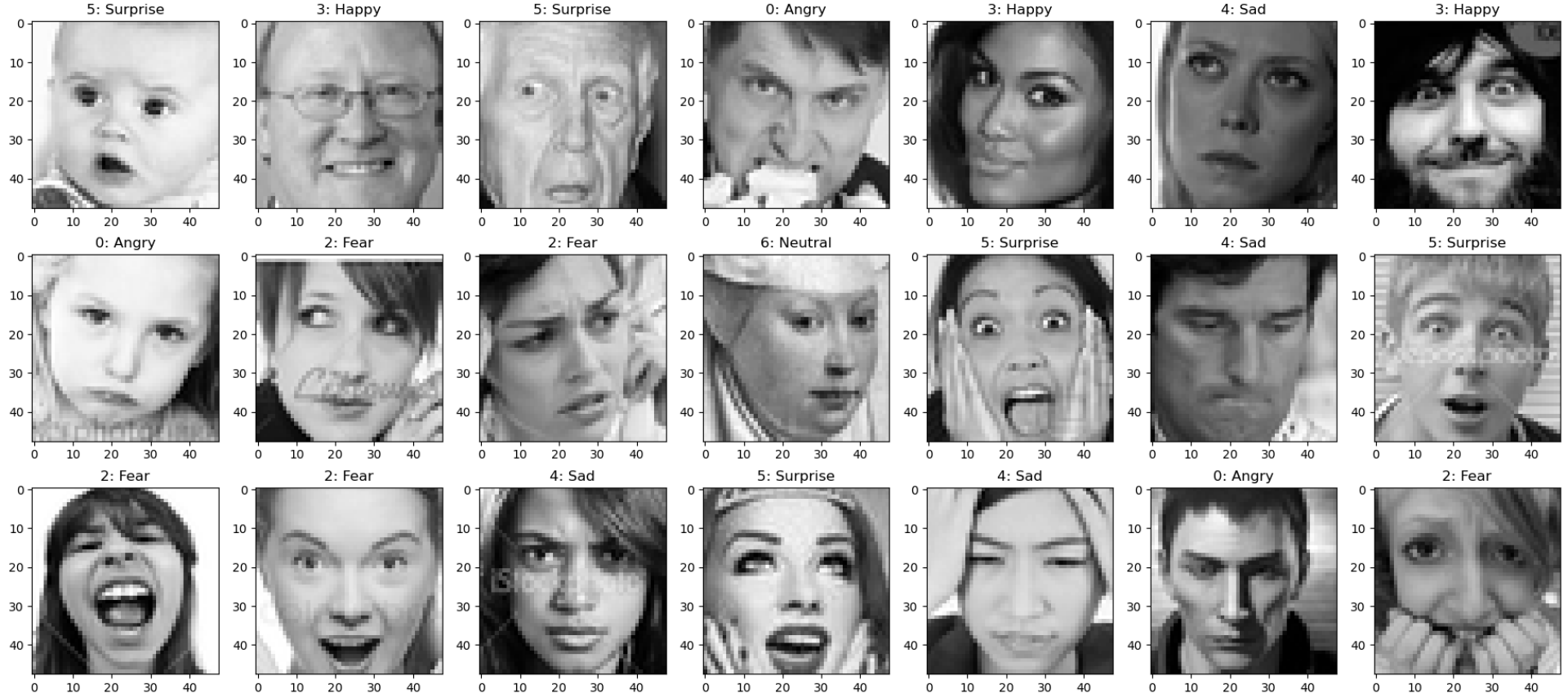
Emotion recognition based on the FER 2013 dataset

- **You should do a model which classifies images into 7 classes:**
  - 'Angry', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprise', 'Neutral'.

- **The training data and labels can be downloaded at: fer2013**

- **We provide a notebook to split the training dataset as well as the public and private test sets.**

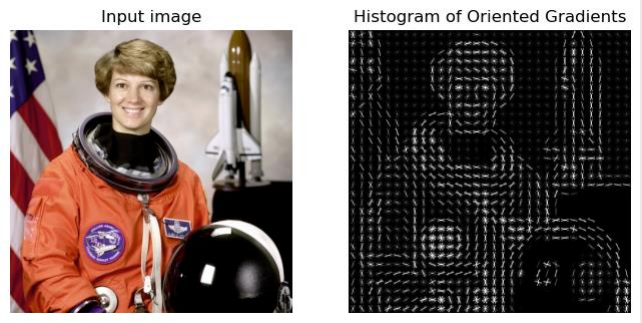
Sample images FER 2013

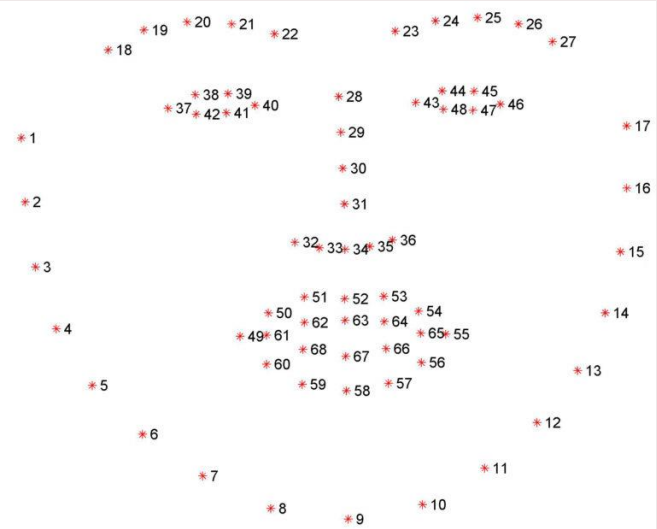


TU/e

What has to be implemented?

- **The feature extraction**
  - HOG features, landmarks, gradients, etc.
  - Data standardization
- **Models**
  - SVM
  - MLP
  - Random forest
  - Fuzzy classifier





What has to be implemented?

- **Evaluation**
  - Accuracy of each model
  - Balanced accuracy
  - Confusion matrix
    - Model should be trained on **Training set,** fine tuned based on the **public test set** and the final performance should be measured on the **private test set.**
    - Report both, the confusion matrix on the public and private test sets.

- **Demo python script that shows the best-performing model operating on real time with the laptop's camera.**

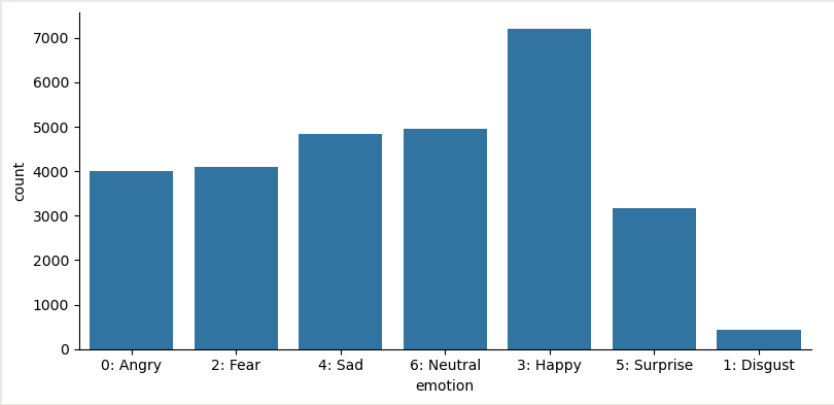
What has to be delivered?

- **Report describing your processing pipelines (feature extraction, data augmentation), the performance of each model (accuracy, confusion matrix).**
- **Code for feature extraction and training each of the reported models**
- **A demo script which uses the laptop camera to load the best of your models and classify images in real time.**


WARNING!

Technical challenges

- **The dataset is imbalanced and large**
- **You have to deal with this by weighting differently each class**
  - Scipy has tools for this as well
  - Augmentations may help as well
- **The feature choice is very important, both in terms of computations and in terms of performance**



Technical challenges (2)

- **The performance achievable in this dataset is not extremely high**
  - Accuracy of 0.5 is already challenging
  - 0.75 achieved with CNNs (which we are not using!)
  - Minimum balanced accuracy of 0.4 **for at lest one of the models measured on the public test set.**
  - **For better performance in your demo script accuracy will be a better representation of the performance of your model.**

Fuzzy classifier

- **Python libraries: PyFUME.**
  - PyFUME (https://github.com/CaroFuchs/pyFUME).
    - Documentation: https://pyfume.readthedocs.io/en/latest/#.
    - Paper: https://ieeexplore.ieee.org/document/9177565.
  - Needs Numpy (Numpy 1.24.4) **HIGHLY** recommended new environment for this!

- **Fuzzy systems work best with low-dimensional features.**
  - For this you may need to use low-dimensional features, to use dimensionality reduction or to use feature selection.

What is expected in the report?

- **The decisions you performed**
  - Which features, standardization augmentations you choose?
  - Why did you choose those, why are these design choices beneficial for this task?
  - Which loss and hyperparameters are chosen for every model.
- **How did you solve the technical challenges**
  - Data augmentation, class weighting, etc.
- **Which challenges did you faced for each specific model implemented and how did you tackled them?**
  - Does the model is better suited for larger or lowe-scale problems?
  - Does the model overfits easily? Do you perform any decision to address that?

What is expected in the report?

- **What is the accuracy, balanced accuracy and confusion for each of the models?**
  - What does these numbers tell us from the dataset and the models?
  - numbers in the public and private test sets agree?
- **Which of the models that you implemented is better suited for this task?**
  - Given the previous considerations which model (and features) works best?
- **Qualitative performance when testing in camera?**
  - Did it perform as good as expected? Yes/no? If not why you think that?