Contents

[Literature Review 2](#_Toc168940095)

[Data Processing 3](#_Toc168940096)

[Data Collection 3](#_Toc168940097)

[FER2013 3](#_Toc168940098)

[Expressions-in-the-Wild (expw) 5](#_Toc168940099)

[Model Validation 8](#_Toc168940100)

[Own Test Data 8](#_Toc168940101)

[References 9](#_Toc168940102)

[Table of Figures 10](#_Toc168940103)

# Literature Review

# Data Processing

## Data Collection

For our facial expression recognition project, we collected two datasets: the FER2013 dataset and the Expressions in the Wild (expw) dataset. The FER-2013 dataset was obtained from a Kaggle challenge (Challenges in Representation Learning: Facial Expression Recognition Challenge | Kaggle, 2013; Goodfellow et al., 2013). The second dataset we found on a website of a research project about learning social relation traits from face images (Zhang et al., 2015a; Zhang et al., 2015b). We stored the datasets in Google Drive and shared the drive amongst the team members for further usage.

## FER2013

We first planned to only use the FER2013 Dataset which consists of 35’887 48x48-pixel grayscale images of faces expressing various emotions. The emotions are categorized into seven emotions as integer values (0=angry, 1=disgust, 2=fear, 3=happy, 4=sad, 5=surprise, 6=neutral). The dataset comes in form of a csv-file and is already split into a training set containing 28’709 images and a testing set containing 7’178 images. It has a column with pixel values and a column with the associated emotions.

For our analysis of the data, we used a jupyter notebook. We loaded the dataset into a pandas dataframe and replaced the integer values of the emotions with the actual string values. After that we took some samples of the dataset and plotted the images with the associated emotions using Python’s Image Library and matplotlib.

A collage of different people's faces

Description automatically generated

Figure 1: A sample of FER2013 images with associated emotion

We thought the images looked clean and were already cropped around the faces. Most of the captions were correct in our opinion. However, there were a few instances where we felt the images could have been labelled differently. We figured that different cultural imprints might have an influence on how we perceive emotions. For this reason, and due to time constraints and limited resources, we decided not to change the labels.

After that we looked at how the labels were distributed. We plotted a bar chart with the label counts for each emotions using plotly express. We saw that the dataset is imbalanced. We have very few images associated with disgust and a lot more images associated with happy.

A graph of different colored bars

Description automatically generated

Figure 2: Distribution of emotions in FER2013 dataset

This imbalance can lead to misclassifications because the model does not have enough data to learn all classes. To tackle that imbalance we thought about looking for another dataset and then combine these two datasets and check the distribution again.

## Expressions-in-the-Wild (expw)

As previously mentioned, we wanted to tackle the imbalance of the FER2013 dataset. We also wanted to find a more diverse dataset with images in various situations. This could help us make our model more robust.

We found the “Expressions in the Wild” dataset which contains a folder with 91’793 colourful images in various sizes and a separate file with the labels. The labels came in the form of a List-file (.lst), which we changed into a csv file.

The labels are as follows:

* image\_name
  + Name of the image file
* face\_id\_in\_image
  + Indicates the specific face within the image. There are images with multiple faces.
* face\_box\_top, face\_box\_left, face\_box\_right, face\_box\_bottom
  + These labels define the bounding box coordinates of the detected face
* face\_box\_confidence
  + Represents the confidence score associated with the face detection
* expression\_label
  + 0: angry, 1: disgust, 2: fear, 3: happy, 4: sad, 5: surprise, 6: neutral

We created a pandas dataframe for the labels and added the headers for each column because the file came without the headers. We then took a sample of 30 items from the dataset using the pandas sample function and a random state of 42. We then used the image name to load the image using the OpenCV library and used the face box coordinates to draw a rectangle around the faces with a margin of 10%. We then plotted the images using matplotlib. We wanted to get a first feel for the data.

A collage of images of people

Description automatically generated

Figure 3: Sample of images with faceboxes before further processing

As you can see the purity of the data is difficult to determine in this form, so that is why we took samples of 30 items with face box confidence of under 50% from the dataset using the same method as before and 30 images with face box confidence of higher than 50%. We then cropped the image around the faces with help of the face box coordinates and plotted the images using matplotlib.

*A collage of people's faces

Description automatically generated*

Figure 4:Sample images with facebox confidence lower than 50%

*A collage of different people

Description automatically generated*

Figure 5: Sample images with facebox confidence higher than 50%

After looking at the samples we found that the images are correctly labelled and look good enough to use for training, even the images with low confidence values. Additionally, they look similar to the FER2013 dataset which is important. If we want to merge both datasets.

A graph of different colored bars

Description automatically generated

Figure 6: Distribution of emotions in Expressions-in-the-Wild dataset

# Model Validation

## Own Test Data

To evaluate our model trained on this data, we are using an AI image generator to create 10 images per expression (total of 70) which will then be annotated by 10 different people.

After that we will use our model on live images from a camera feed and see how well it can detect facial expressions.

# References

*Challenges in Representation Learning: Facial Expression Recognition Challenge | Kaggle*. (2013). Retrieved February 29, 2024, from https://www.kaggle.com/competitions/challenges-in-representation-learning-facial-expression-recognition-challenge/data

Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D., Zhou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shawe-Taylor, J., Milakov, M., Park, J., . . . Bengio, Y. (2013, July 1). *Challenges in Representation Learning: A report on three machine learning contests*. arXiv.org. https://arxiv.org/abs/1307.0414

Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2015a). *Learning Social Relation Traits from Face Images*. Retrieved April 12, 2024, from https://mmlab.ie.cuhk.edu.hk/projects/socialrelation/index.html

Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2015b, September 14). *Learning Social Relation Traits from Face Images*. arXiv.org. https://arxiv.org/abs/1509.03936

# Table of Figures

[Figure 1: A sample of FER2013 images with associated emotion 3](#_Toc168939634)

[Figure 2: Distribution of emotions in FER2013 dataset 4](#_Toc168939635)

[Figure 3: Sample of images with faceboxes before further processing 5](#_Toc168939636)

[Figure 4:Sample images with facebox confidence lower than 50% 6](#_Toc168939637)

[Figure 5: Sample images with facebox confidence higher than 50% 6](#_Toc168939638)

[Figure 6: Distribution of emotions in Expressions-in-the-Wild dataset 7](#_Toc168939639)