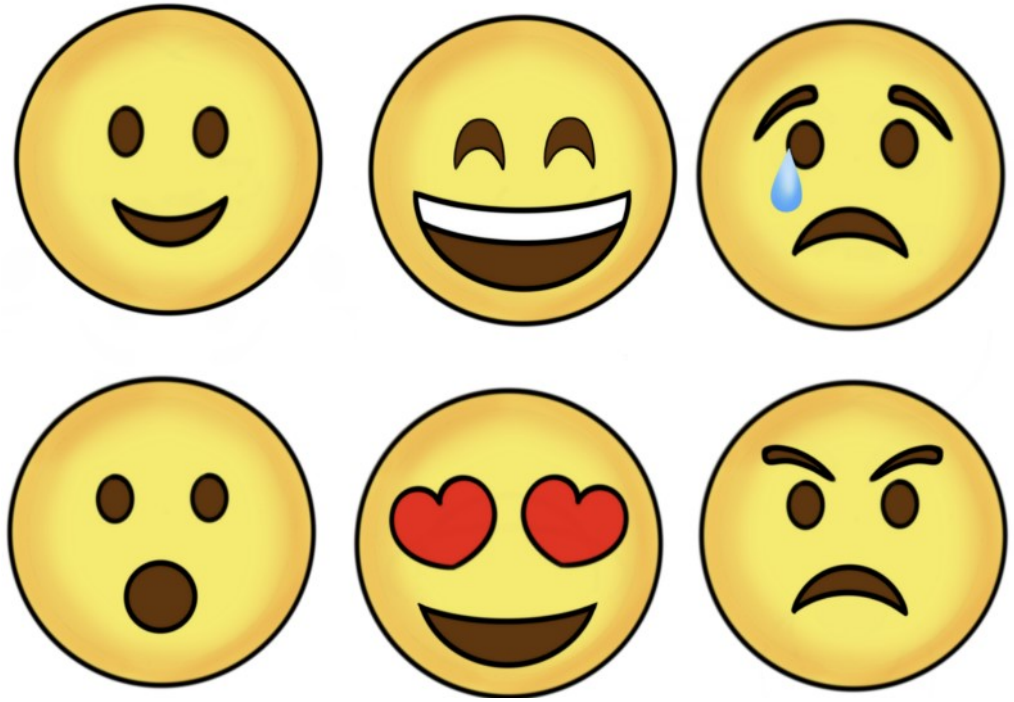
FACIAL EXPRESSION DETECTION USING DEEP LEARNING



David von Holzen & Maitane Iturrate-Garcia

Module 6: Deep Learning – CAS in Applied Data Science

MOTIVATION

The ability to recognize and name emotions and expressions is crucial for:

- creating social connections and interactions
- experiencing empathy and sympathy

(Source: https://www.helpcenterapp.com/blog/empathy-in-customer-service/)

Neurotypical children:

- basic facial expression recognition at 3-4 months
- responding to others' emotions by 7 months

(Source: Helga Annenkova (Shutterstock))

MOTIVATION

One of the main diagnostic characteristics of **Autism Spectrum Disorder (ASD):** the difficulty of recognizing and understanding emotions, including challenges in identifying facial expressions and alexithymia (inability to recognize and label the emotions they feel).

Real-time **facial expression recognition (FER)** apps may help people with ASD recognizing emotions.

https://link.springer.com/article/10.1007/s11042-022-13558-9

PROJECT OVERVIEW – AIM

To classify accurately basic facial expressions (e.g., angry, fear, happy, sad, confused, neutral, surprised) using convolutional neural networks (CNNs).

Because of computational limitations, only happy vs. angry classification task in Python (`tensorflow`).

PROJECT OVERVIEW – DATASET

Dataset: FER-2013

48 × 48 pixel grayscale images

Faces more or less centred

About same amount of space in each image

Public test set: 3,589 examples

Training set: 28,709 examples

7 face expression categories from 0 to 6:

angry (0), disgust (1), fear (2), **happy (3)**,

sad (4), surprise (5), neutral (6)

https://www.kaggle.com/datasets/msambare/fer2013/data

PROJECT OVERVIEW – PIPELINE

1. **INITIALISATION**
   (loading libraries, setting directories, data loading…)
2. **DATA PRE-PROCESSING**
   (data generator, data augmentation, data split)
3. **MODEL DEFINITION**
   (classes, input shape, layers: convolutional/ hidden/output, flattening)
4. **MODEL TRAINING**
5. **MODEL VALIDATION**
6. **REPORTING**

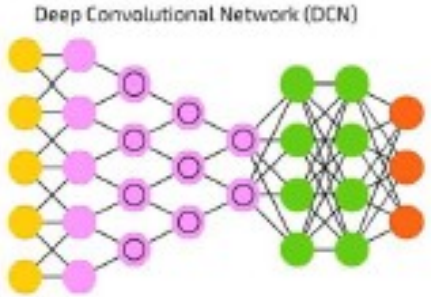
Data pre-processing

- data generator (on-the-fly batch of images generation during training)
- data augmentation (to increase the size and diversity of the training set:
  - rescale pixel values to be between 0 and 1
  - randomly shift the width and height of images by up to 10 %
  - flip images horizontally at random
  - zoom in on the image
  - shift the image counter-clockwise by the indicated degrees
- split data into training and validation data

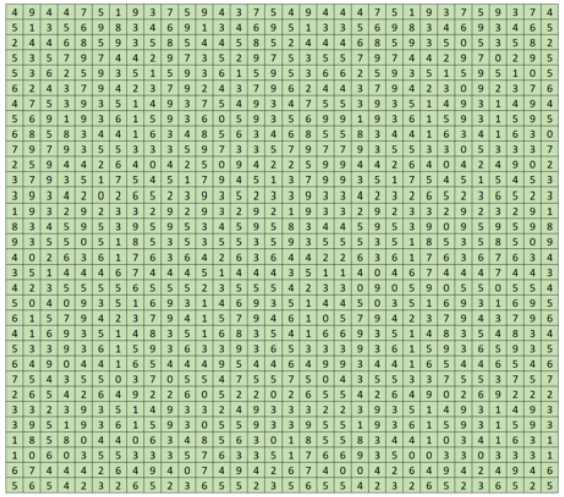
PROJECT OVERVIEW – PIPELINE

1. **INITIALISATION**
   (loading libraries, setting directories, data loading…)
2. **DATA PRE-PROCESSING**
   (data generator, data augmentation, data split)
3. **MODEL DEFINITION**
   (classes, input shape, layers: convolutional/ hidden/output, flattening)
4. **MODEL TRAINING**
5. **MODEL VALIDATION**
6. **REPORTING**

Model definition



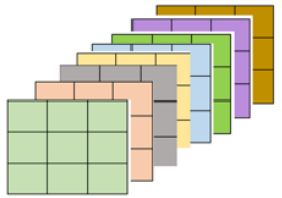
Input



32 × 32 × 1

CONV+POOL layer

Filters (kernel)



3 × 3 × 8

Conv2D,
MaxPool2D

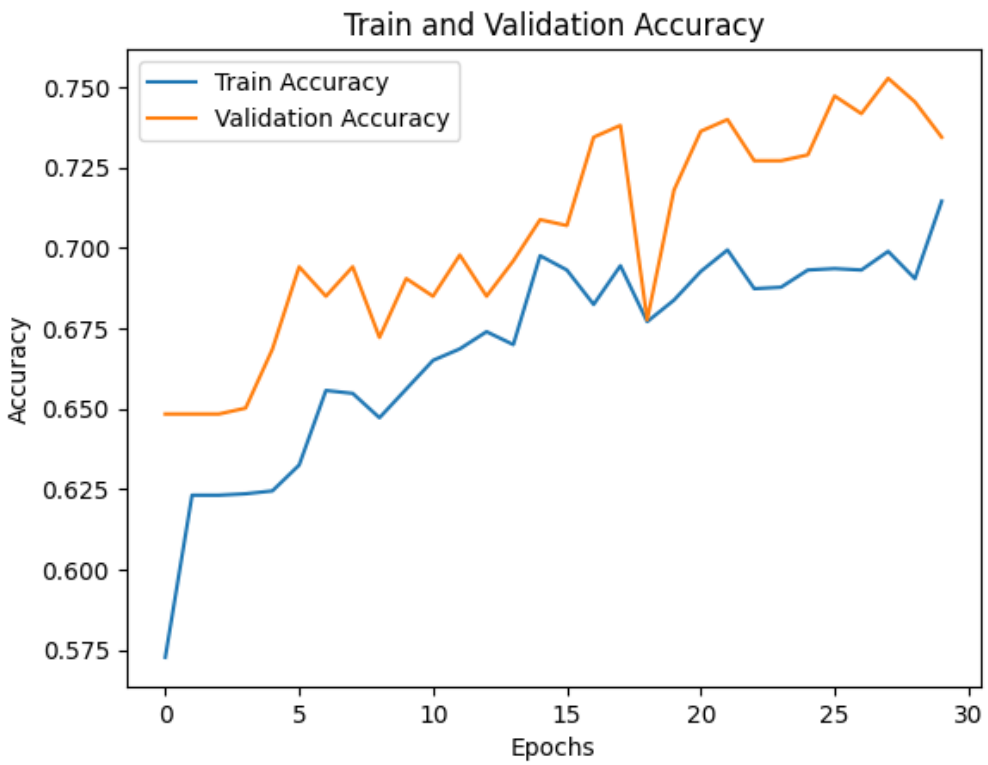
FC (ReLU)

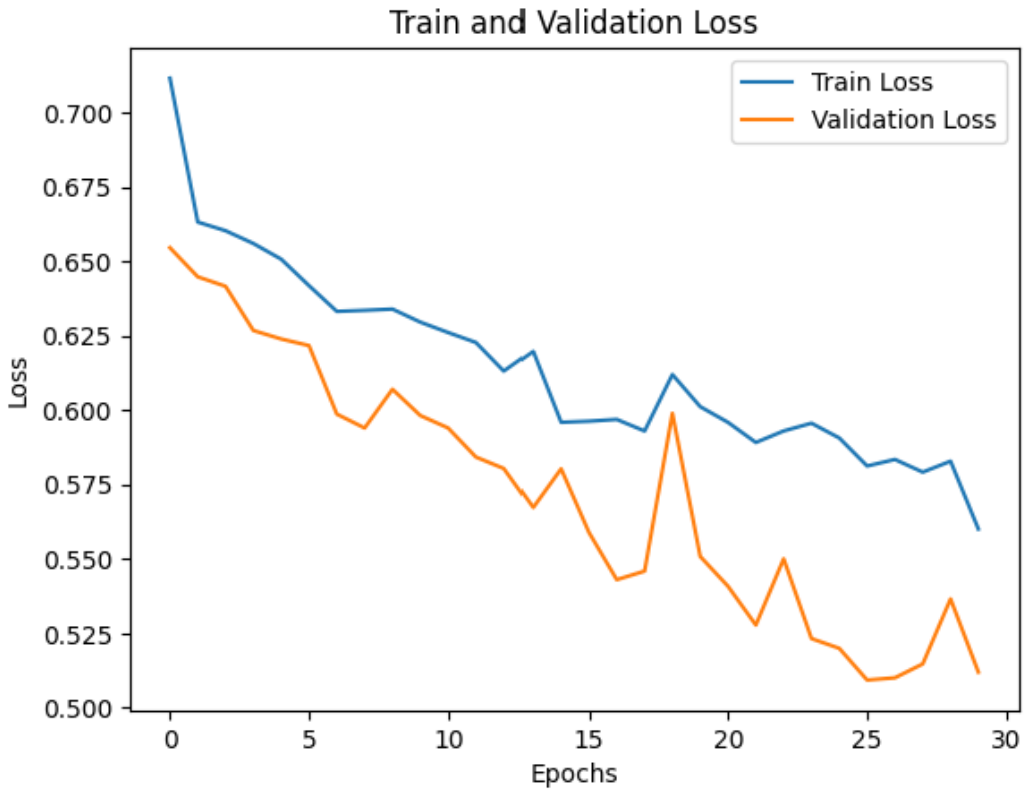
Output (sigmoid)

0 = Angry

1 = Happy

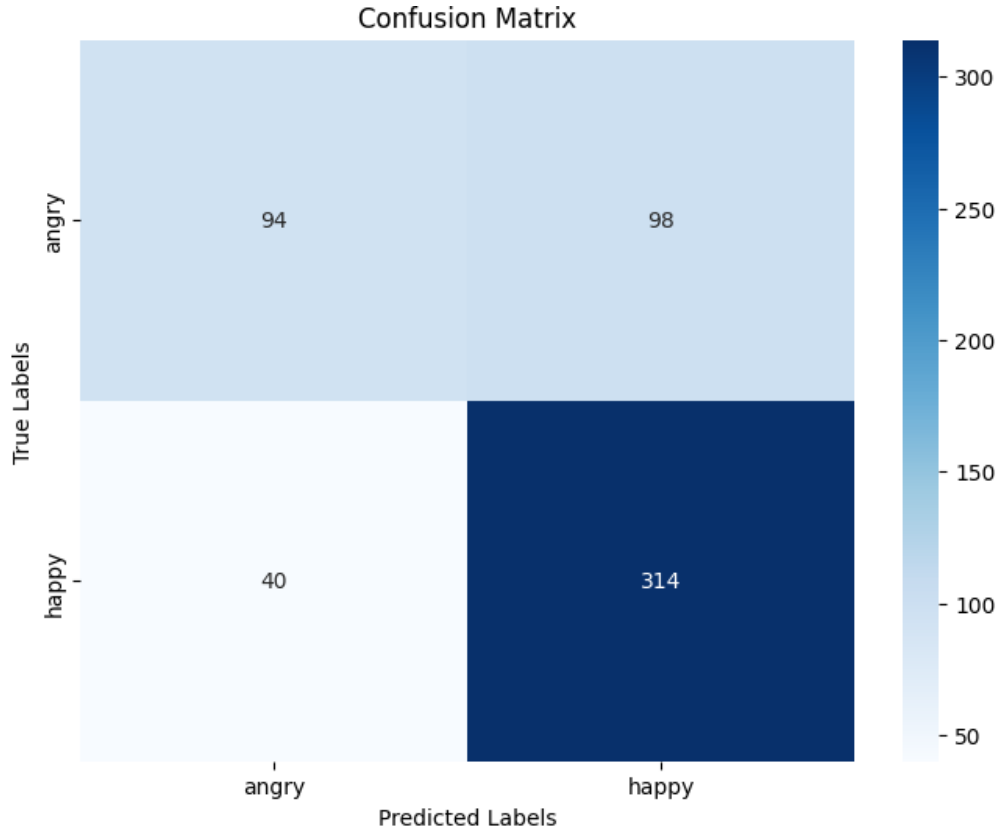
ACCURACY & LOSS



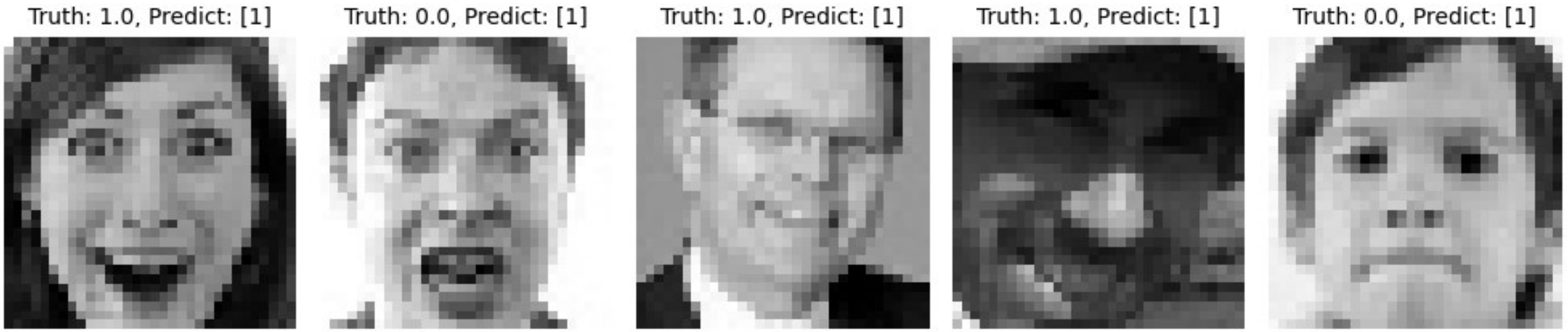


CLASSIFICATION REPORT & CONFUSION MATRIX

	precision	recall	f1-score	support
angry	0.70	0.49	0.58	192
happy	0.76	0.89	0.82	354
accuracy			0.75	546
macro avg	0.73	0.69	0.70	546
weighted avg	0.74	0.75	0.73	546



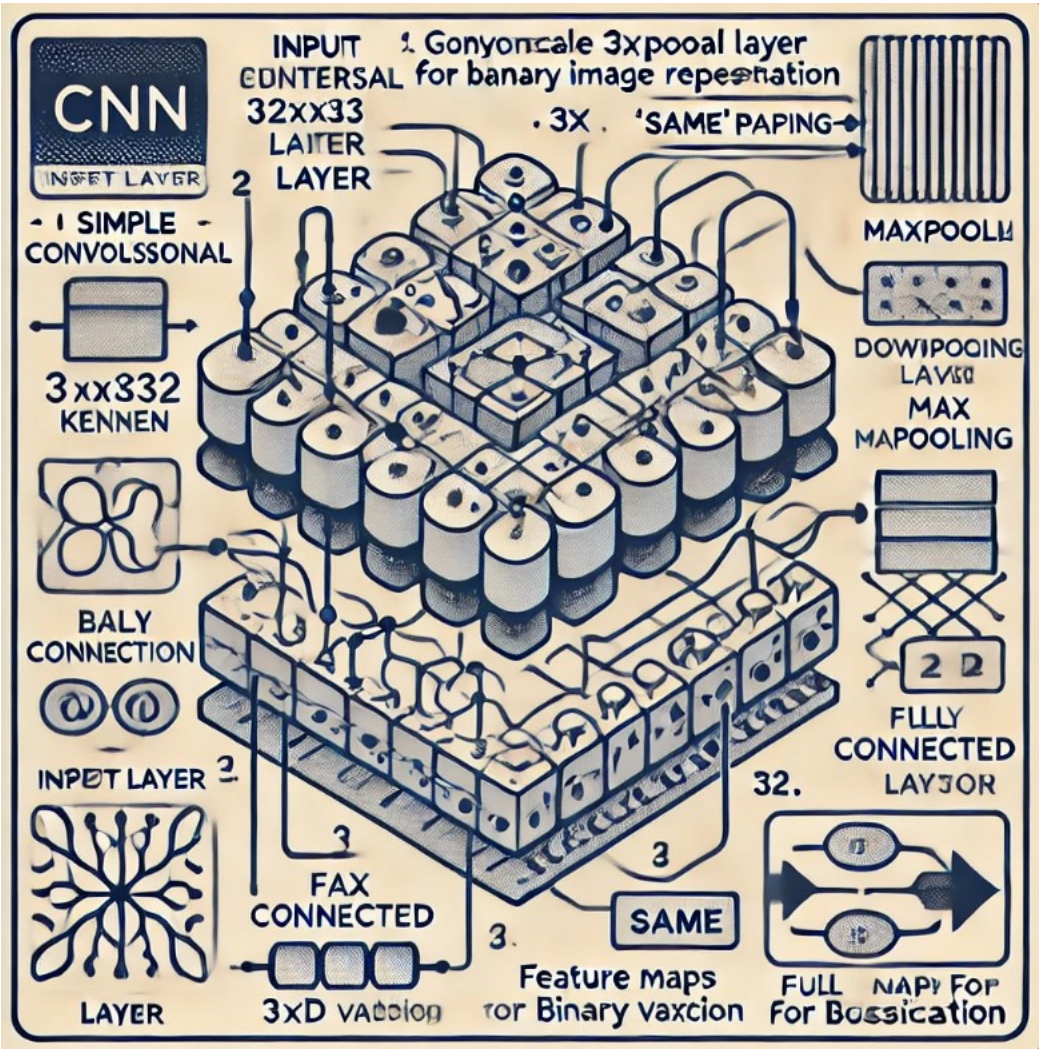
VISUALIZING PREDICTIONS



MODEL OPTIMIZATION AND FURTHER STEPS

1. Modify the parameters of the image data generator (e.g. by reducing zoom, shear or rotation range)
2. Increase the batch size
3. Use another dataset with images of better quality and balanced number (i.e. same number of images in each category of feelings)
4. Adapt and use pretrained models (e.g. EfficientNetV2)
5. Add more convolutional layers with more filters for a more complex model
6. Use other feelings with greater contrast (e.g. happy vs. sad)
7. Move from a binary classification to multiple classification (e.g., happy, sad, surprise, angry)

DISCUSSION


CNN
INPUT
Gonyoncale 3xpooal layer
for banary image repesnation
32xx33
LAITER
LAYER
. 3X .
'SAME' PAPING
MAXPOOLI
- I SIMPLE -
CONVOLSSONAL
DOWIPOOING
LAVSO
3 xx332
KENNEN
MAX
MAPOOLING
BALY
CONNECTION
FLLLY
CONNECTED
LAYJOR
INPDT LAYER
32.
FAX
CONNECTED
SAME
LAYER
3xD
Feature maps
for Binary vaxcion
FULL
For Bossication