FACIAL EMOTION DETECTION

Using Machine Learning

PROJECT TEAM

PATRICK BACZA

ALLISON WELLS

WILLIAM BROWN

CHRIS WINN

SYED HUSAIN

AGENDA

01

FER Overview

02

Approach and model description

03

Results

04

Real-time demo

WHAT IS FACIAL EMOTION RECOGNITION?

- Technology that analyzes facial expressions to help determine emotional state
- Uses static images and/or videos
- Goal to automate the process of determining emotions in real-time
- Analyzes various facial features (eyebrows, eyes, mouth, etc.)
- Decoding emotion expressions an interest for decades in psychology
- AI/deep learning utilized to automate and increase accuracy
- CNN ideally suited for this effort
- 55% of emotional information is conveyed visually (Mehrabian, 1967)

POSSIBLE USES OF EMOTION RECOGNITION

- Check emotions at live performances
- Real-time customer satisfaction
- Product testing
- Market research- focus groups
- Political debate/town hall audience reaction
- Education- curriculum adaptation
- Mental health diagnosis
  - Employee safety
  - Patient care
  - Automotive safety
- Video game- adapt to player mood
- Fraud detection
- Virtual personal assistants/robots
- Employee interviews

CONCERNS

- Data Privacy
  - Data breaches
  - Transparency and informed consent
  - Inference of special categories of personal data (health data, political opinions)
  - Accessibility of data by non-authorized entities (employers, health insurers, etc.)
- Data accuracy
  - Multiple emotions experienced simultaneously
  - Camera angles, lighting
  - FER usually treated as 100% accurate
- Fairness
  - Racial/cultural bias
  - Medical conditions causing misclassification
- Media and academic criticism
  - “Microsoft is removing emotion recognition features from its facial recognition tech”, NBCNews.com, June 27, 2022
  - “Artificial Intelligence Is Misreading Human Emotion”, The Atlantic, April 27, 2021

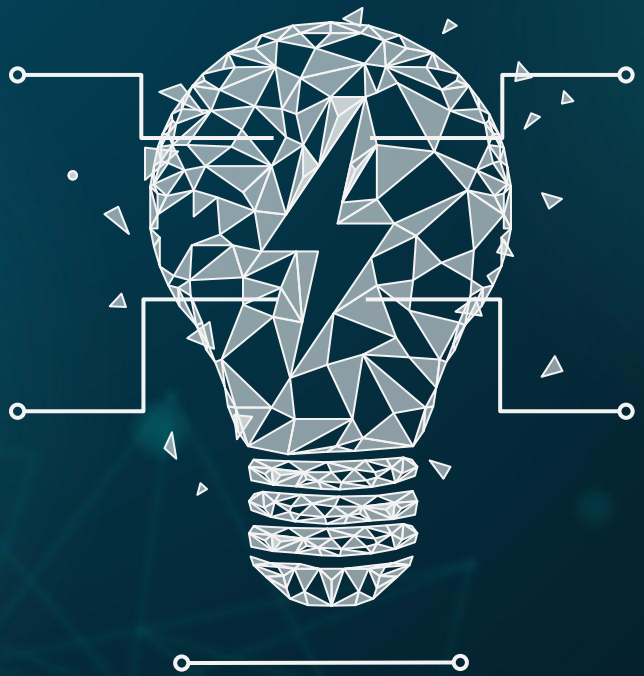
FER2013 CONTEST

CONTEST

Associated with ICML 2013
Workshop on
Representative Learning

PRIZE MONEY

First Prize- $350
Second Prize- $150
(generously donated by Google)



OF ENTRIES

56 teams
Only 4 beat “null” model
(CNN with no learning
except in final classification
layer)

WINNING ACCURACY

71.162%

FER2013 DATASET

GREYSCALE IMAGES OF FACES

48 x 48 pixels

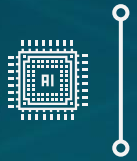
7 CATEGORIES OF EMOTIONS

Angry | Fear | Disgust | Happy | Sad | Surprise | Neutral

ROBUST DATASET

35,888 faces

RESOURCES USED



AWS/S3

TensorFlow/Keras

Pandas

SKLearn

APPROACH

- Preprocess data
- Create model
- Evaluate model
- Make predictions

Preprocessing

Create arrays for each image

Separate out the emotion assigned to each image

Split data into the test and training data sets

Apply Data Augmentation to the data

Data Augmentation

- Method used to alter images in the dataset ex. rotate, zoom in, flip it, increase width , increase height.
- Benefits:
  - Increases the size of the data set .
  - Exposes model to a wider range of variation in the data which makes the model stronger and better to handle real world scenarios where images will have all different types of orientations and sizes.
  - Reduces overfitting - Since random variations are added to the data, the model is prevented from relying too heavily on specific details or relationships within the data.
  - Can address class imbalance.
  - Overall, the model will be much better at making predictions on new unseen data.

Data Augmentation

Rotation range = 20 - rotates images randomly within a range of -20 to 20 degrees

Width shift range = 0.2 - shifts width by up to 20% of total width

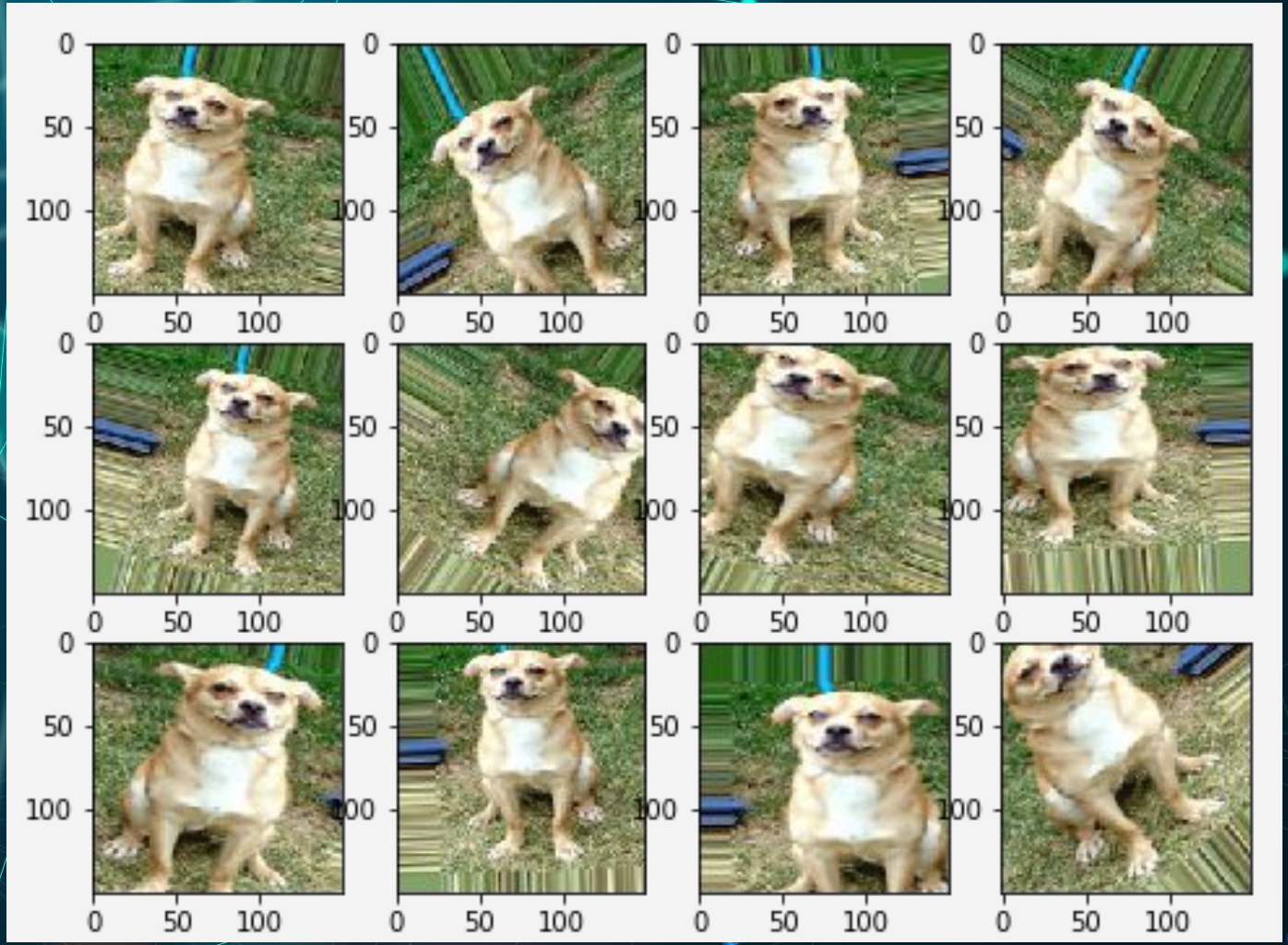
Height shift range = 0.2- shifts height by up to 20% of total height

Shear range = 0.2 - Changes the viewpoint of the image with range of -0.2 to 0.2

Zoom_range = 0.2 - zooms in images by up to 20%

Horizontal flip = True - Mirrors image

Fill mode = nearest - Fills in any empty pixels with the values of the nearest pixel


0
50
100
0 50 100

Layer Definitions

Convolutional Layer - A filter or window that slides over the input data, capturing local patterns or features.

Dense layer - a layer where every neuron is connected to every neuron in the previous layer. It performs a linear transformation followed by a nonlinear activation allowing the network to learn and represent complex patterns in the data.

Batch normalization - Adjusts values of the inputs within each mini-batch of training data to ensure that they have similar statistics, such as mean and variance. Helps the network learn more efficiently

Dropout Layer - Randomly deactivates a fraction of the neurons in a layer during training to introduce noise and prevent overfitting. It forces the model to learn more independent features

Maxpooling layer - Divides the input data into non-overlapping regions and selects the maximum value from each region. By selecting the maximum value from each region, it captures the strongest activation or presence of a feature in that region.

Creating the model

At first, the model did really well on the training, but not on testing.
Overfitting

Approaches to overfitting:

- Drop out layers
- Data Augmentation
- Adding more learning layers (dense layers)
- Slower learn rate

Creating the model

INITIAL

Increasing drop-out layers and experimenting

LEARNING

Added dense layers to increase learning

OVERFITTING

Used data augmentation to combat

SIMPLIFY

Removed dense layers and decreased drop-out

TESTING

Real time application with webcam view.

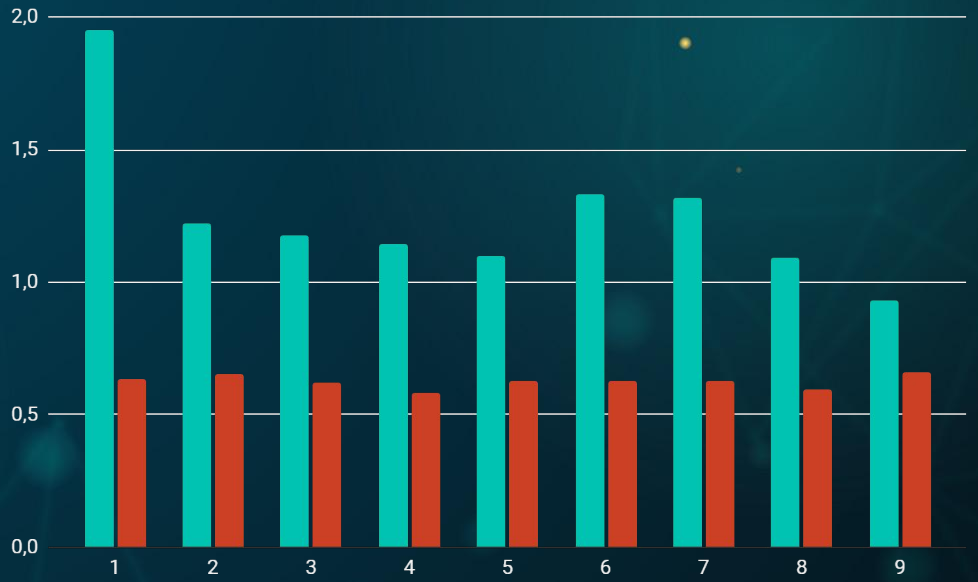
Final Model

- 10 Convolutional layers
  - Between 64 and 512 neurons for each layer
- 8 batch normalization layers
- 4 max pooling layers
- 3 dropout layers
  - Each deactivation rate is set at 10%
- 1 dense layer
  - Output layer
  -
- Training Accuracy: 69.7%
- Training loss: 0.82
- Testing Accuracy: 66.5%
- Testing Loss: 0.93

RESULTS

LOSS

ACCURACY



DEMO

Real-time FER