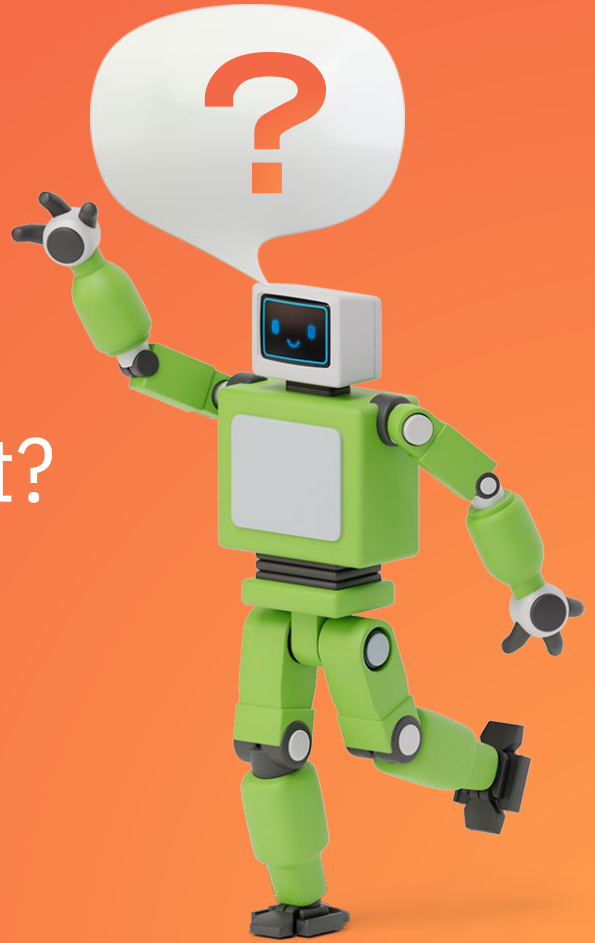
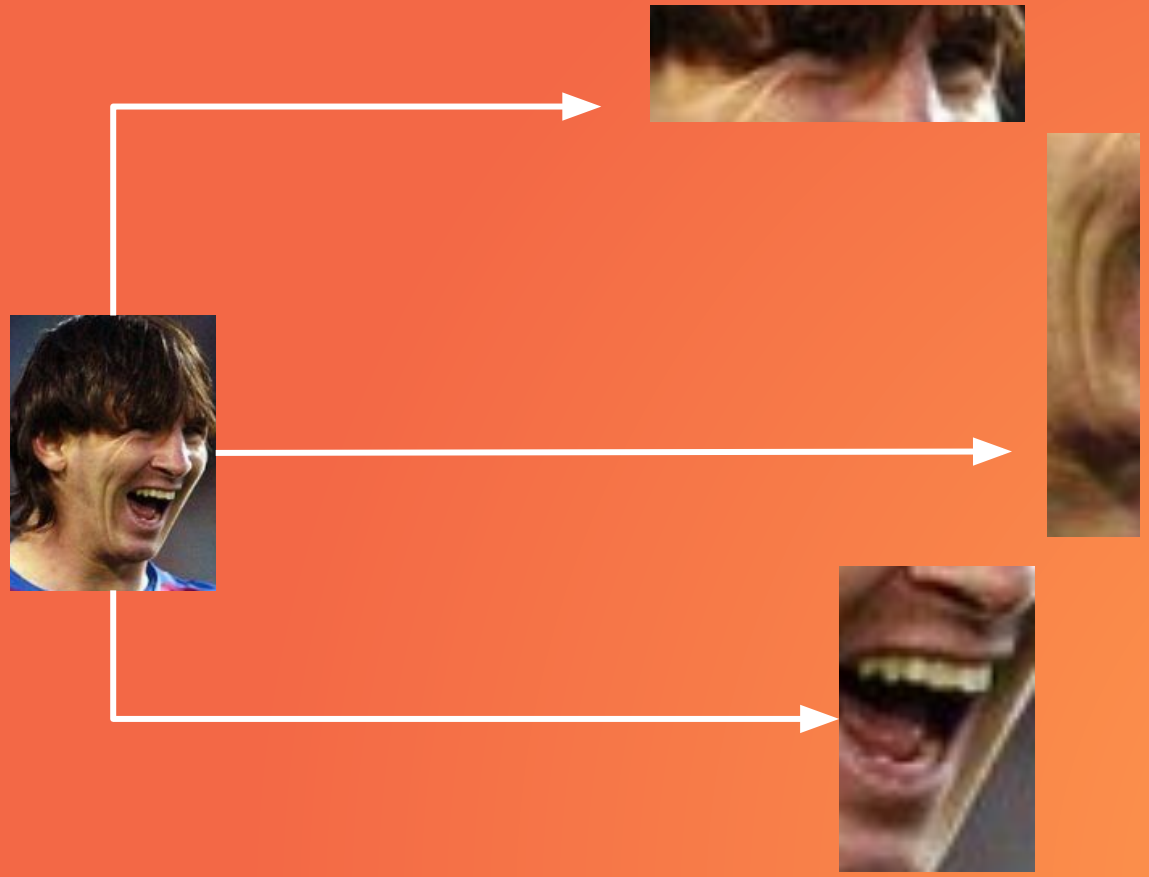
FACIAL EXPRESSION RECOGNITION (FER)

APS360H-2022W Team 55
Litao (John) Zhou - 1006013092
Jiping (Peter) Li - 1005983269
Yang (Jack) Chen - 1005747649

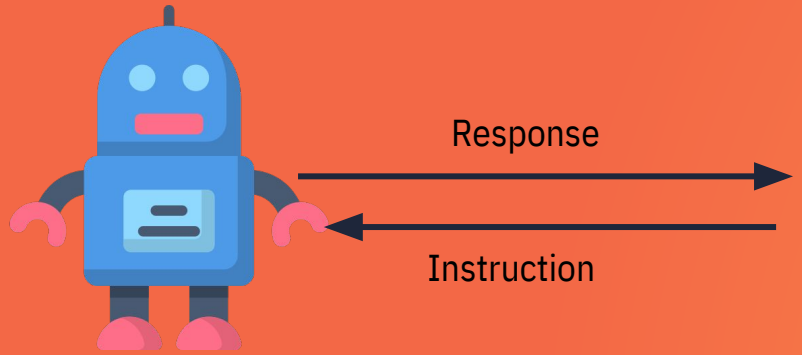
But....How do you know that?



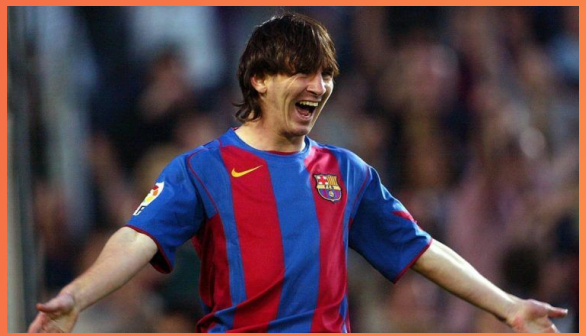
From the detail of face…

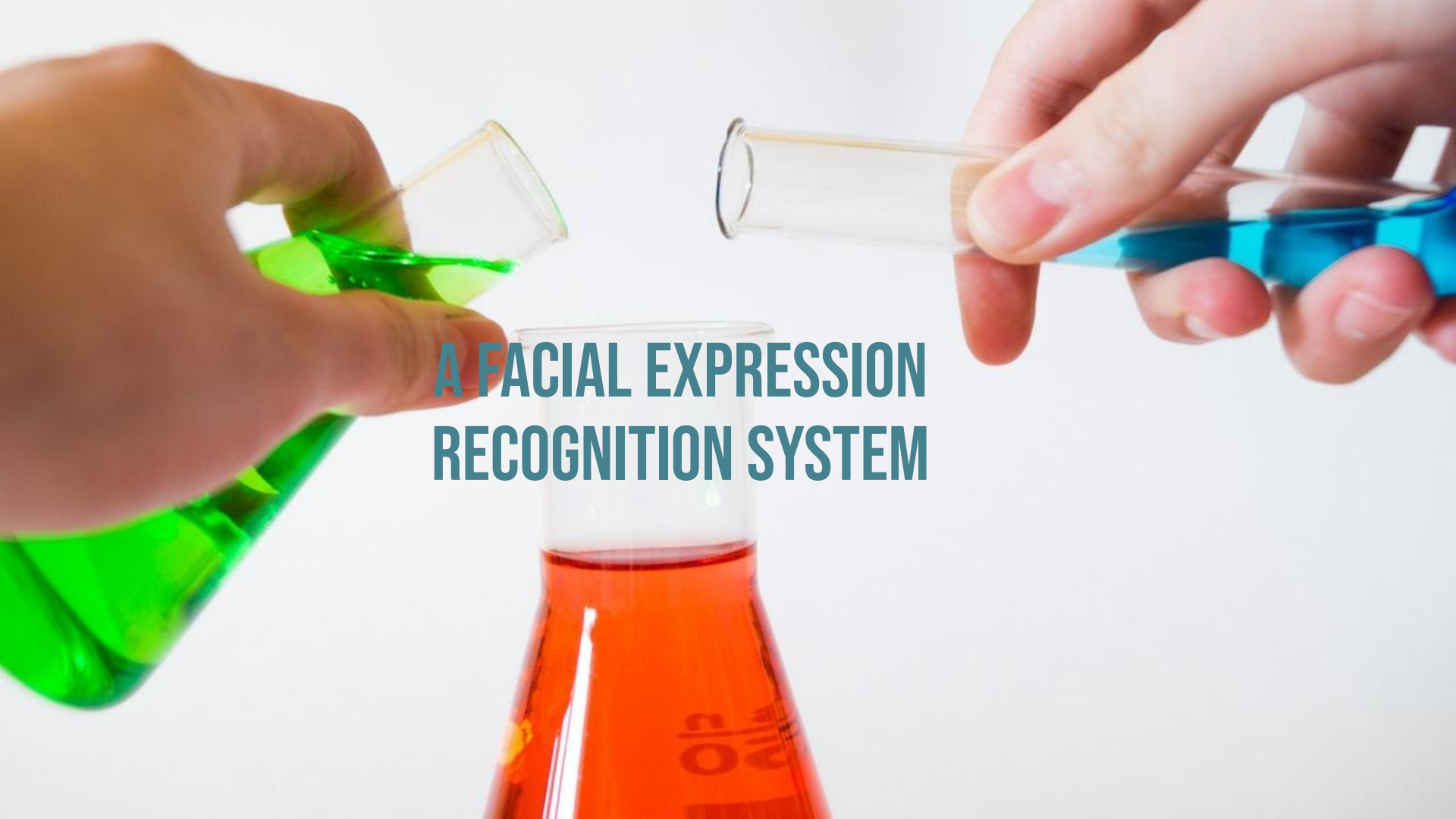


Facial expression recognition is not hard for human while for computer that is not the case…



Computer




A FACIAL EXPRESSION
RECOGNITION SYSTEM

DATA PROCESSING

How we process our data

DATASET

100000 IMAGES, 18 EXPRESSIONS

EXAMPLE : CRYING

SINGLE FACE

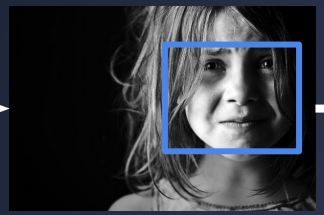
MULTIPLE FACES

NO FACE

FLOWCHART

ONLINE DATASET

LOCALIZATION



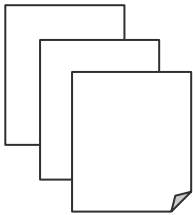
CLASSIFICATION

“POSITIVE”

“NEUTRAL”

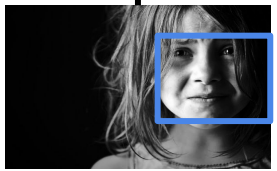
“NEGATIVE”

LOCALIZATION MODEL

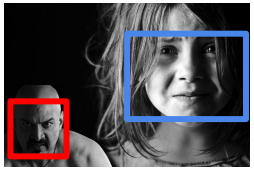


ONLINE DATASET

GRAYSCALE



SINGLE FACE



MULTIPLE FACES

CHOOSE LARGEST

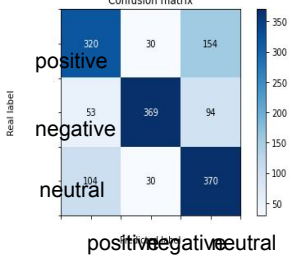
CROP

RESIZE

LOCALIZATION

(PRE-TRAINED CASCADE CLASSIFIER)

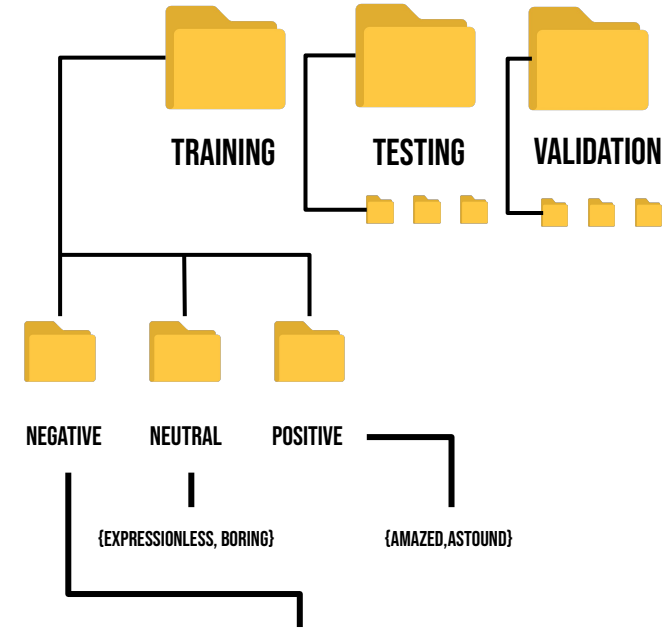
FIND NO HUMAN FACE, DISCARD IMAGE

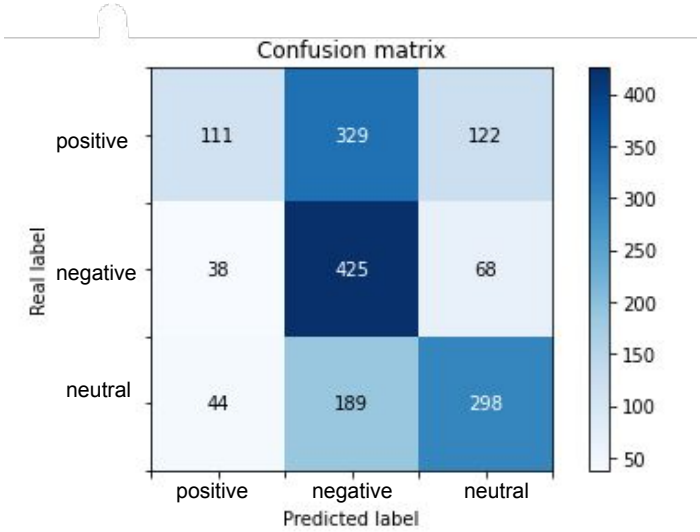


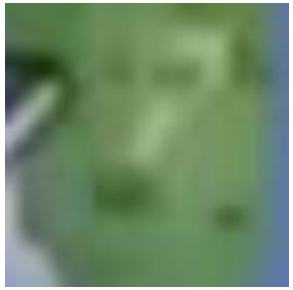
$$\textit{Precision (positive)} = \frac{\textit{True Positive}}{\textit{Total Predicted Positive}}$$

$$\textit{Recall (positive)} = \frac{\textit{True Positive}}{\textit{Total Actual Positive}}$$

$$\textit{F1 (positive)} = 2 \times \frac{\textit{Precision} * \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$




Confusion matrix
positive
negative
neutral
Real label
111
329
122
38
425
68
44
189
298
positive
negative
neutral
Predicted label
400
350
300
250
200
150
100
50

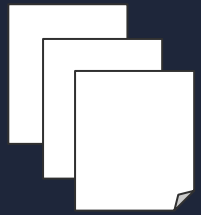



N
O

MUL
TIPL

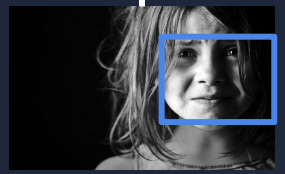
SI
NG

LOCALIZATION MODEL



ONLINE DATASET

GRAYSCALE



SINGLE FACE

MULTIPLE FACES

CHOOSE LARGEST

CROP

RESIZE

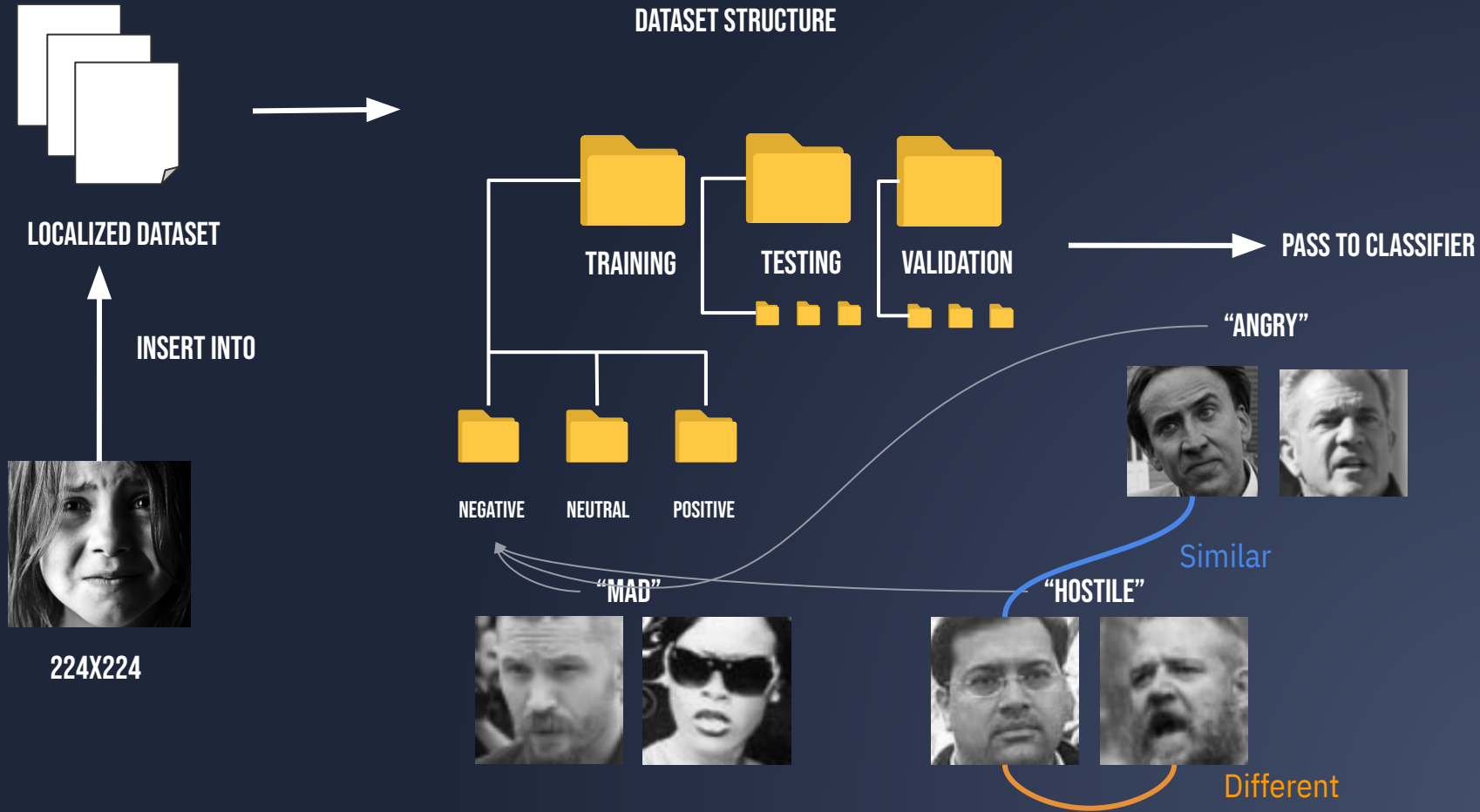
224X224

LOCALIZATION

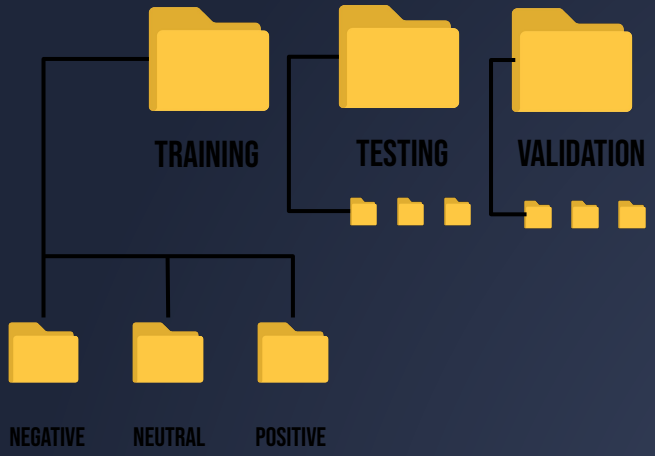
(PRE-TRAINED CASCADE CLASSIFIER)

FIND NO HUMAN FACE, DISCARD IMAGE

DATASET STRUCTURE



DATASET STRUCTURE



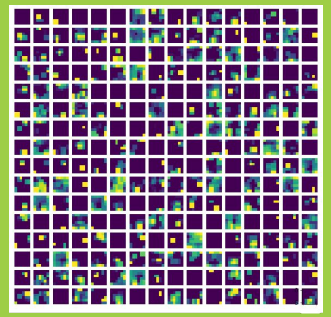
NEURAL NETWORKS

ALEXNET

CNN

CLASSIFICATION MODEL

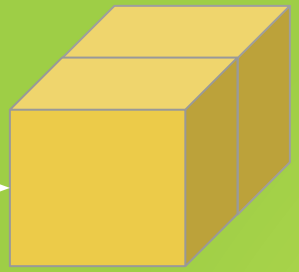
ALEXNET.FEATURES



256X6X6 FEATURES

CONV1

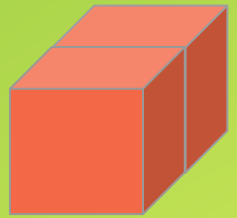
OUT-CHANNELS:200
STRIDE=1
KERNEL=3



200X4X4

CONV2

OUT-CHANNELS:150
STRIDE=1
KERNEL=2



150X3X3

FLATTEN

150X3X3 = 1350

FC1

200

FC2

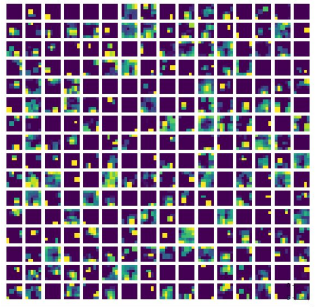
3

POSITIVE

NEGATIVE

NEUTRAL

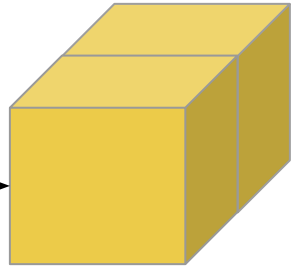
ALEXNET.FEATURES



256X6X6 FEATURES

CONV1

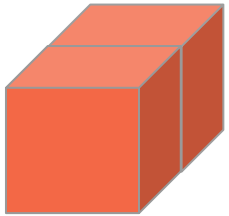
OUT-CHANNELS:200
STRIDE=1
KERNEL=3



200X4X4

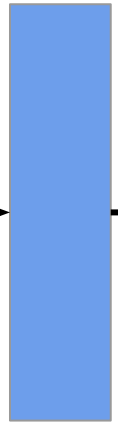
CONV2

OUT-CHANNELS:150
STRIDE=1
KERNEL=2



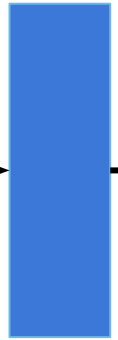
150X3X3

FLATTEN



150X3X3 = 1350

FC1



200

FC2

3

POSITIVE

NEGATIVE

NEUTRAL

CLASSIFICATION MODEL

- **MODEL PARAMS**

  BATCH SIZE : 256

  NUM EPOCHS : 50

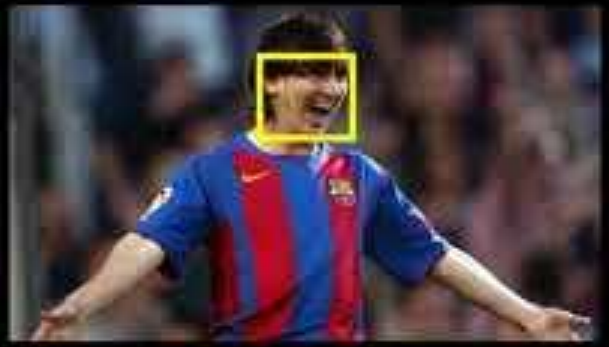
  LEARN_RATE : 0.001

  LOSS FUNCTION : Cross Entropy Loss

  Activation function : Rectified Linear Unit (ReLU)

DEMONSTRATION

Please see the video clip!



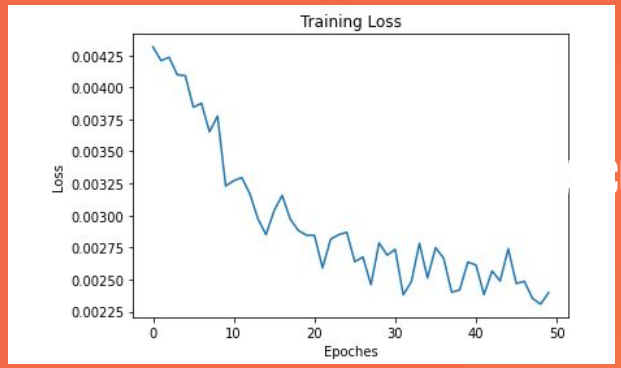
OpenCv

RESULT

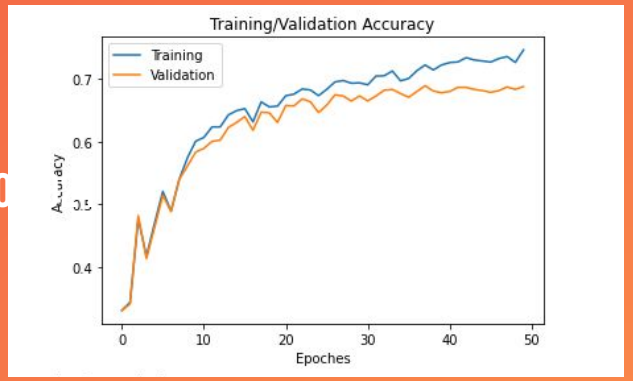
The qualitative and quantitative result of our model

QUANTITATIVE RESULT - OVERALL ACCURACY

RESULT :



URACY : 0

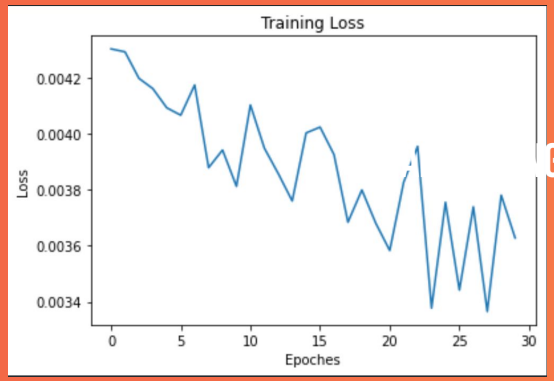


FINAL TRAINING ACCURACY LOSS : 0.00242

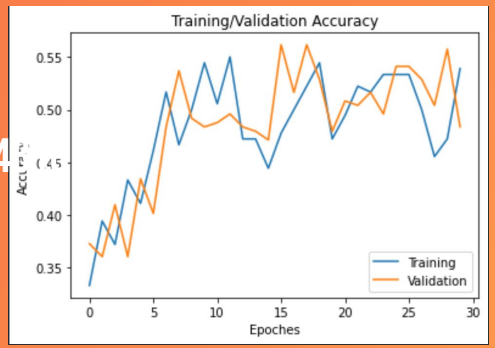
FINAL TRAINING ACCURACY : 0.7458
FINAL VALIDATION ACCURACY : 0.6876

V 26%

BASELINE MODEL:



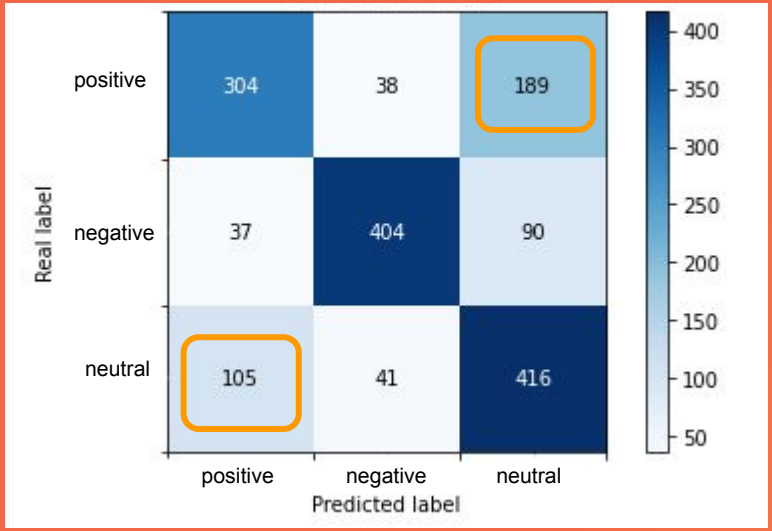
G ACCURACY : 0.4



FINAL TRAINING ACCURACY LOSS : 0.00362

FINAL TRAINING ACCURACY : 0.5388
FINAL VALIDATION ACCURACY : 0.4836

QUALITATIVE RESULT : CONFUSION MATRIX



Not good at differentiate between “neutral” and “positive”

QUALITATIVE RESULT

POSITIVE

POSITIVE NEGATIVE NEUTRAL

[0.4138, 0.0756, 0.5106]

PREDS : NEUTRAL

POSITIVE

POSITIVE NEGATIVE NEUTRAL

[0.1569, 0.1318, 0.7113]

PREDS : NEUTRAL

INSIGHT

The future development path for our model and related improvement

IMPROVEMENT TOWARD OUR CURRENT MODEL

- Limitations of CNN, and might need to test to choose other models for better accuracy
- Better Dataset with less errors in the sample

THANK YOU!