



# COMPUTER VISION

# EMOTION EXPRESSION

# BAG OF VISUAL WORD

# AND DEEP LEARNING

Prepared by:

**DO Nhu Tai (176680)**

May, 16<sup>th</sup>, 2018



# Agenda

## 1. Introduction

## 2. Problem 1: Emotion Bag-of-Word Object

## 3. Problem 2: Emotion Deep Learning

## 4. Problem 3: Emotion Video Extraction

## 5. Future Works: Bag-of-Word using Deep Learning

## Demo Program



# 1. INTRODUCTION

## 1.1. PROBLEMS

- Emotion Expression Recognition: FER2013

Training Images





# 1. INTRODUCTION

## 1.2. DATASET

Validating Images

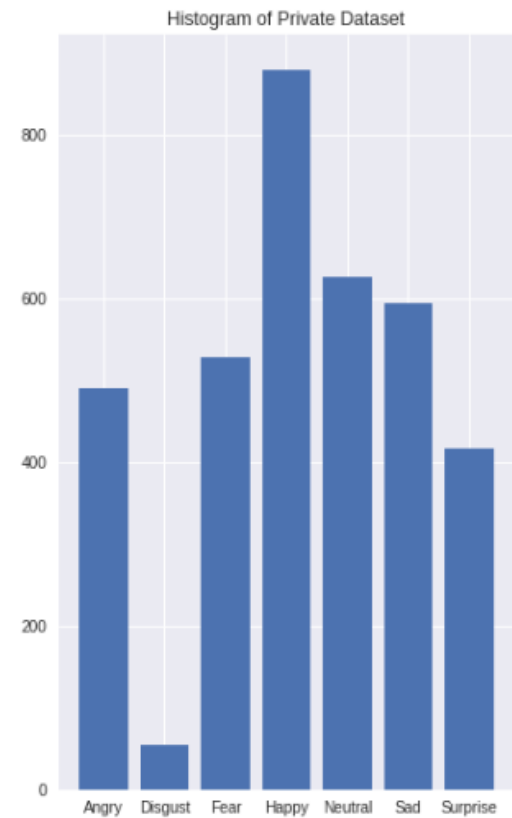
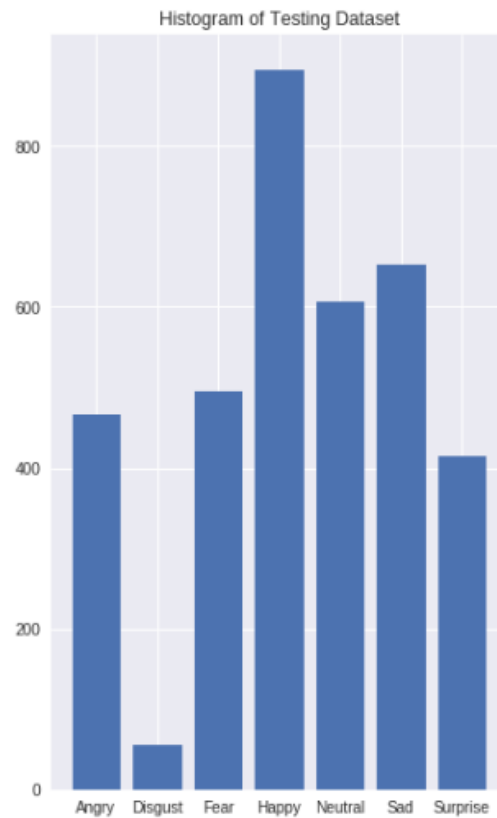
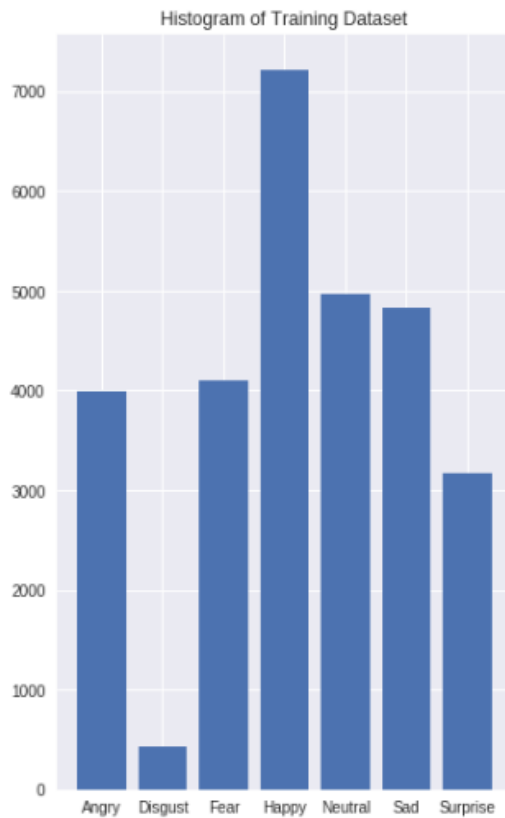




# 1. INTRODUCTION

## 1.3. DISTRIBUTION DATASET

Number of images in the training dataset: 28709  
Number of images in the validation dataset: 3589  
Number of images in the testing dataset: 3589  
Image information: 48 x 48 x 1





## 2. Problem 01: Emotion with Bag-of-Word And Sparse SIFT Featur

### 2.1. Training Process

- Extract descriptors using SIFT
- Merge descriptors into local patches
- Cluster local patches using Mini Batch K-Means to build codewords
- Build Feature Histogram Model based on codewords
- Normalize Feature Histogram Model
- Classify by Multi-Class SVM



## 2. Problem 01: Emotion with Bag-of-Word And Sparse SIFT Featur

### 2.2. Testing Process:

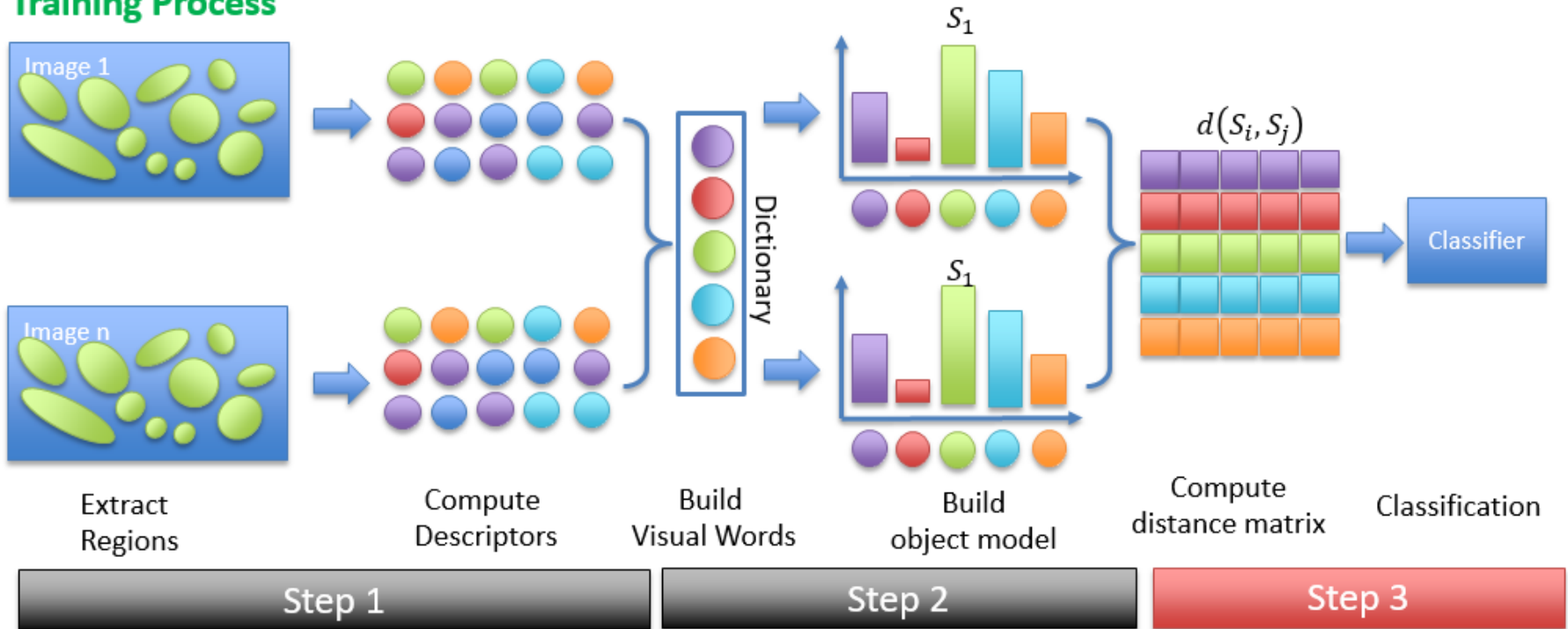
- Extract descriptors using SIFT
- Match descriptors with K-Means Cluster Center to build codewords for testing image
- Create and normalize Feature Histogram Model for Image Codewords
- Predict by Multi-Class SVM



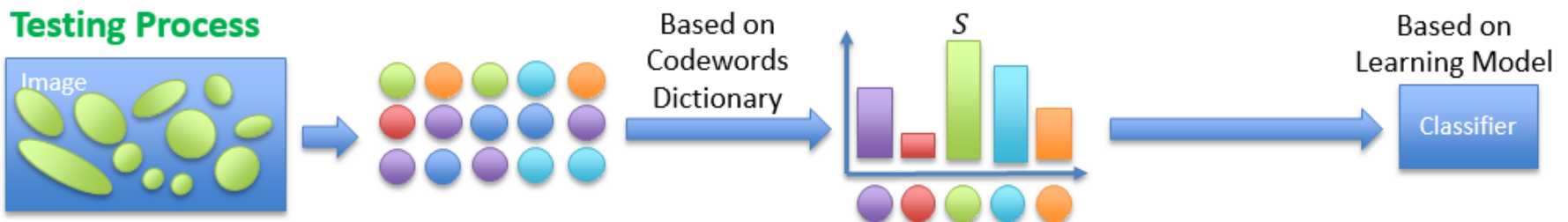
## 2. Prolem 01: Emotion with Bag-of-Word And Sparse SIFT Feature

### 2.3. Process:

#### Training Process



#### Testing Process

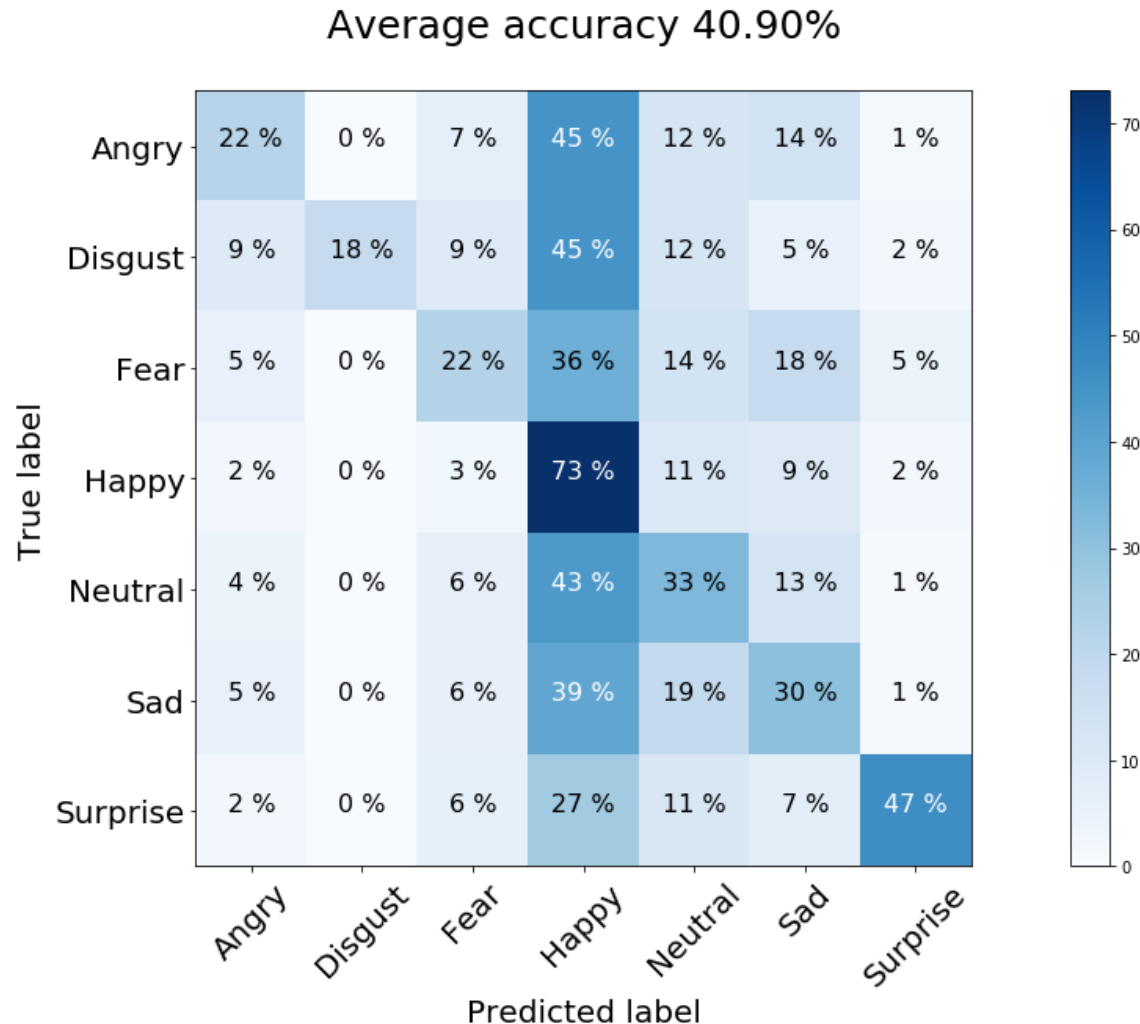






## 2. Prolem 01: Emotion with Bag-of-Word And Sparse SIFT Featur

### 2.4. Results:





## 3. Prolem 02: Emotion with Deep Learning

### 3.1. Deep Learning CNN-Like-VGG16 Model:

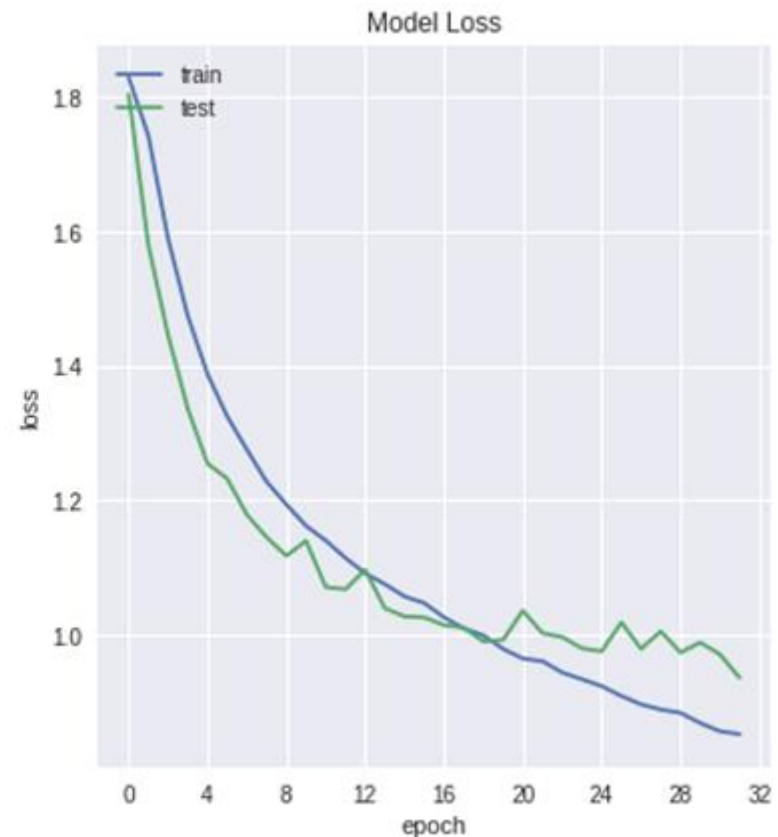
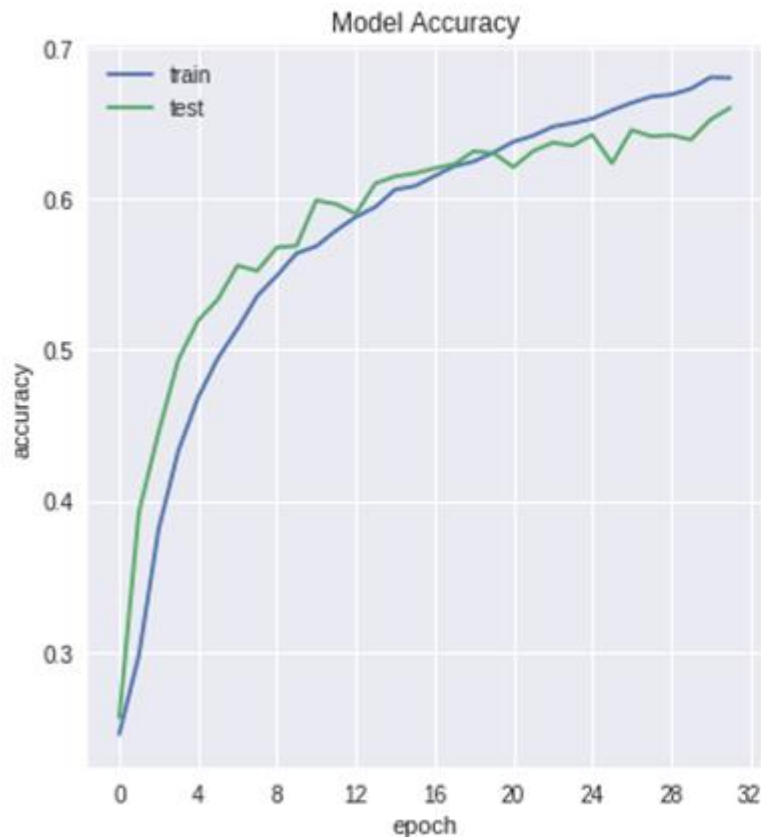
- Block 1 – 3 Conv2D (64, (3,3)), MaxPooling (2,2), Dropout (0.2)
- Block 2 – 4 Conv2D (128, (3,3)), MaxPooling (2,2), Dropout (0.2)
- Block 3 – 4 Conv2D (256, (3,3)), MaxPooling (2,2), Dropout (0.2)
- Classifier – Flatten, Dense (1024), Dropout(0.5), Dense(7, SoftMax)



## 3. Problem 02: Emotion with Deep Learning

### 3.2. Training History

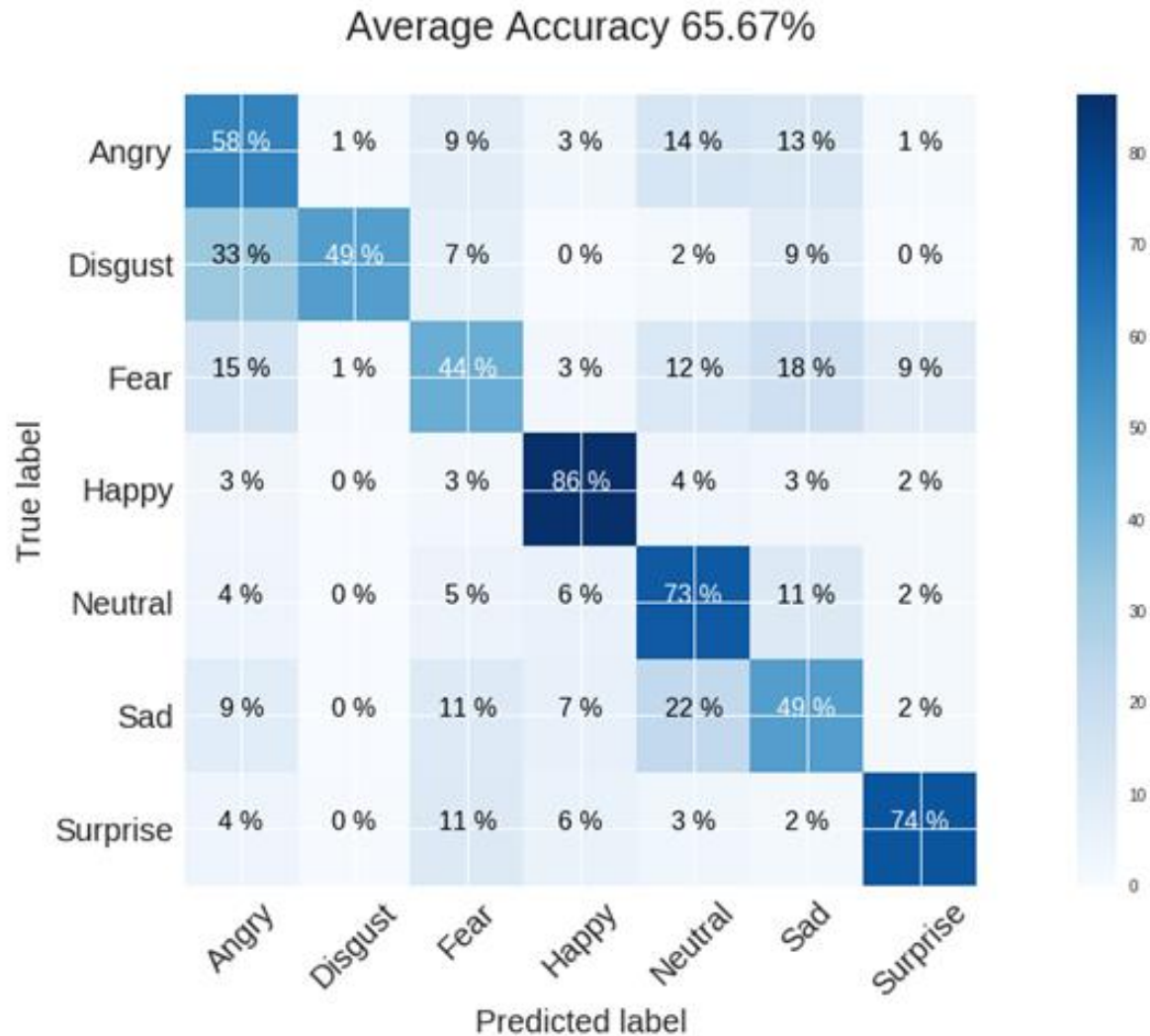
Training at Epoch 31





## 3. Problem 02: Emotion with Deep Learning

### 3.3. Results:





## 4. Prolem 03: Video Emotion Extraction

### 4.1. Program Description:

- Extract small video clips in a video and vote video emotion of small clips

The screenshot shows a video player window titled "Video" with the following data displayed in red text at the top: `FPS: 9 - 5.73 (s) - 1 Face(s)[ 5.30-0.43][2,0,0]neutral`. A green rectangular box is drawn around a person's face in the video frame, with a red label `[[1 (neutral)]]` overlaid on it. Blue arrows point from various labels to specific parts of the interface:

- Face Detection**: Points to the green bounding box around the face.
- Start Track**: Points to the first value in the time range, `5.30`.
- Time**: Points to the second value in the time range, `0.43`.
- Num Split**: Points to the third value in the array, `2`.
- Number of frame From No Face Detect**: Points to the fourth value in the array, `0`.
- Number of frame from Lost Focus Face**: Points to the fifth value in the array, `0`.
- Emotion Voting**: Points to the word `neutral` at the end of the data string.
- Red: Face Focus Green: No Focus**: Points to the red label `[[1 (neutral)]]`.
- Identity Face**: Points to the person's face in the video frame.



## 4. Prolem 03: Video Emotion Extraction

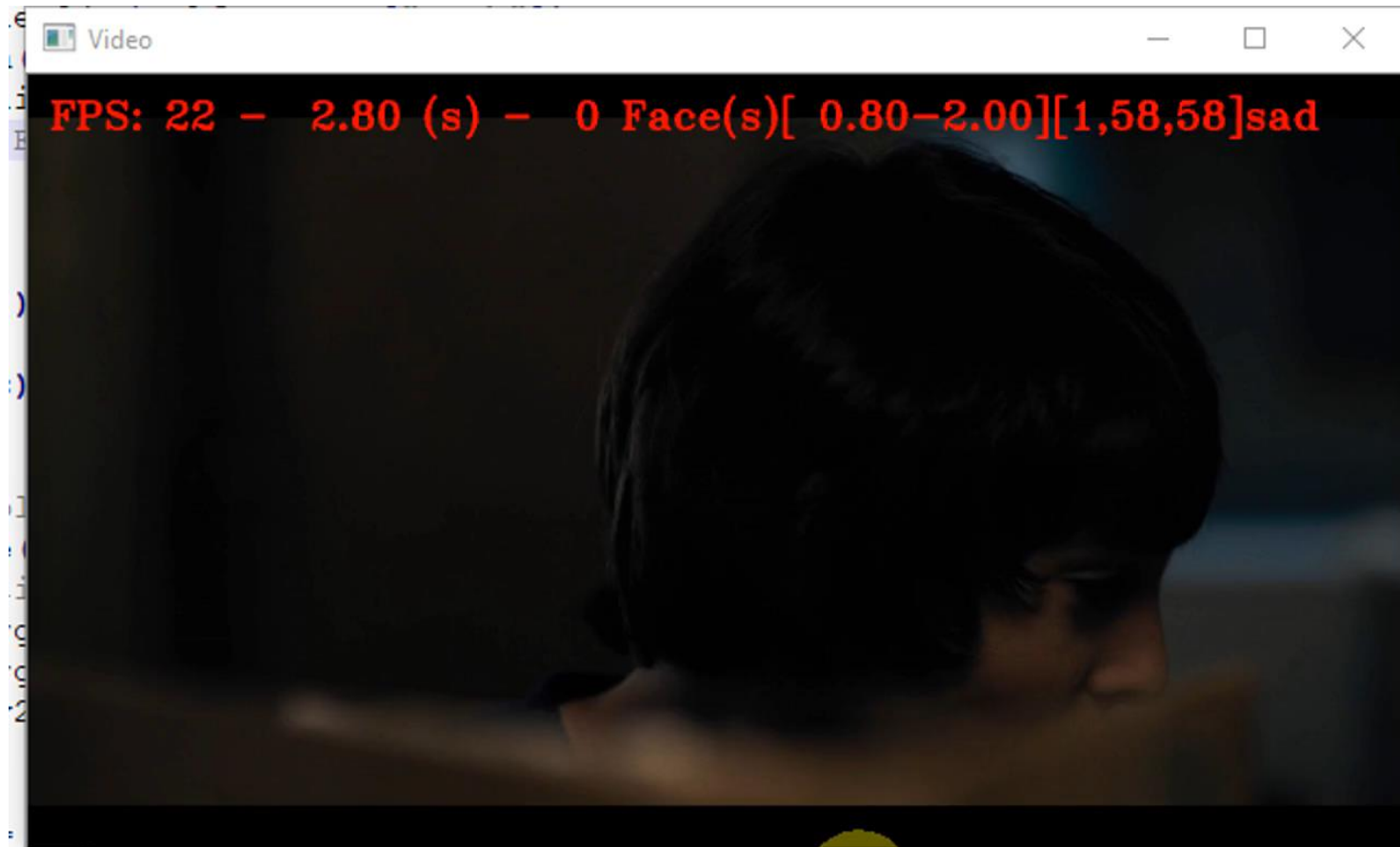
### 4.2. Program Features:

- Console Program with Modules
- Emotion MLCNN,
- Face Detection Dlib, OpenCV, MTCNN,
- Face Matching using Hungarian Method,
- Face Description with VGG Face



## 4. Prolem 03: Video Emotion Extraction

### 4.3. Demo:





## 5. Future Works

### Bag-of-Word using Deep Learning

- Use CNN Model for feature extraction (old: SIFT)
- Use RNN for building histogram model (old: K-Means)





**THANKS FOR LISTENING!**  
**Waiting for question!**