**COMPUTER VISION FINAL PROJECT**

Chonnam National University, Gwangju, South Korea

[2018-06] Computer Vision Class

Professor: Chilwoo Lee

Student: Do Nhu Tai

Target: Emotional Expression with Bag-of-Word and Deep Learning

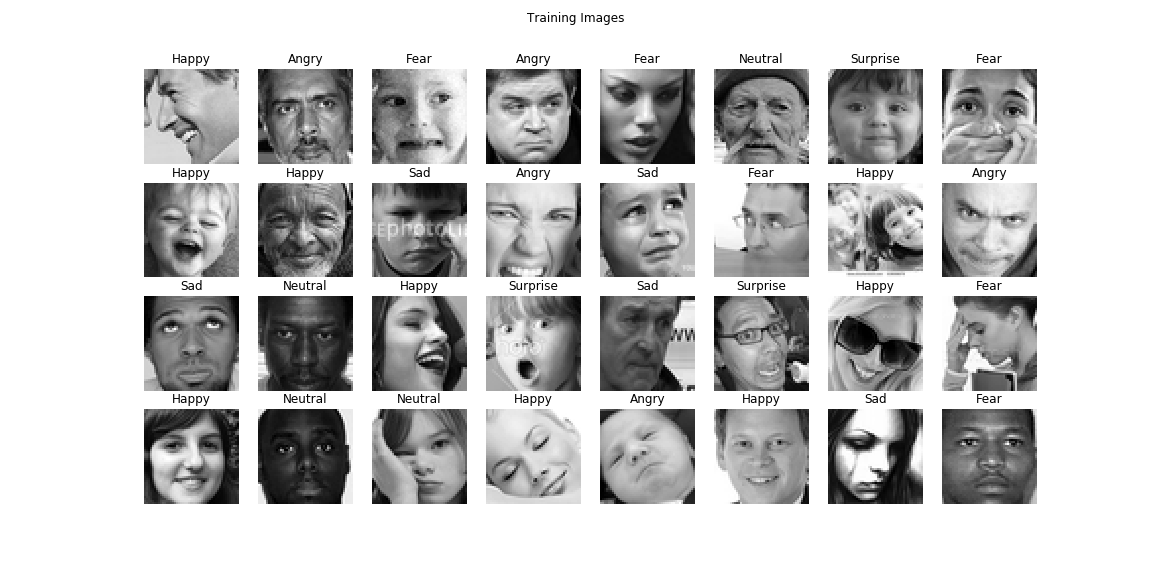
Dataset: [Kaggle Fer2013] - Facial Expression Recognition Challenge

<https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/>

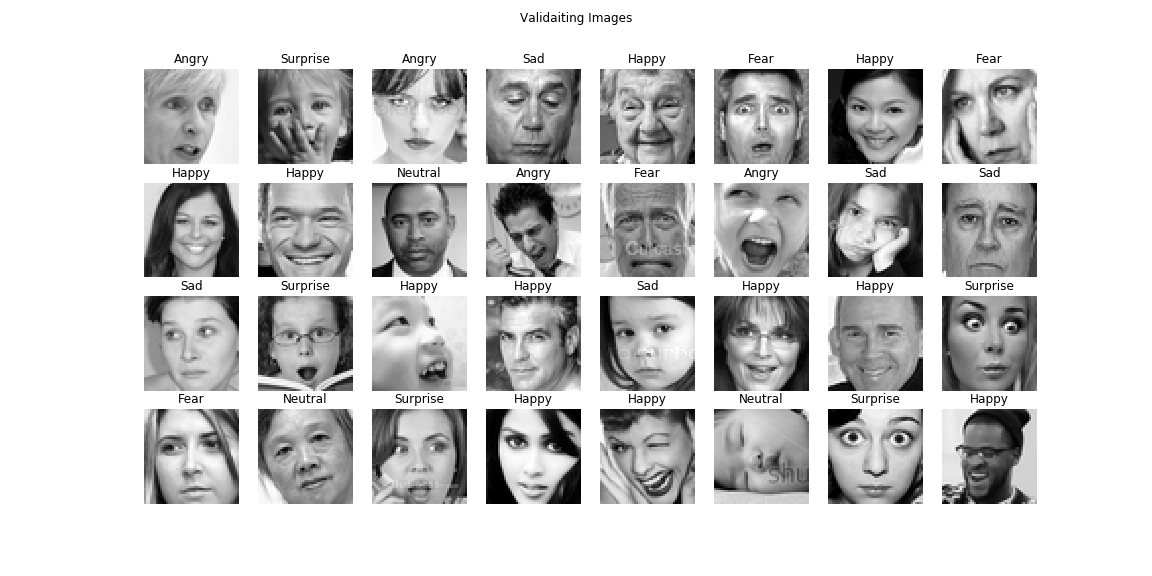
(1) Dataset Information:

Number of images in the training dataset:         28,709  
Number of images in the public dataset:                 3,589  
Number of images in the private dataset:         3,589  
Image information: 48 x 48 x 1

(2) Training Images:

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/training_images.png)

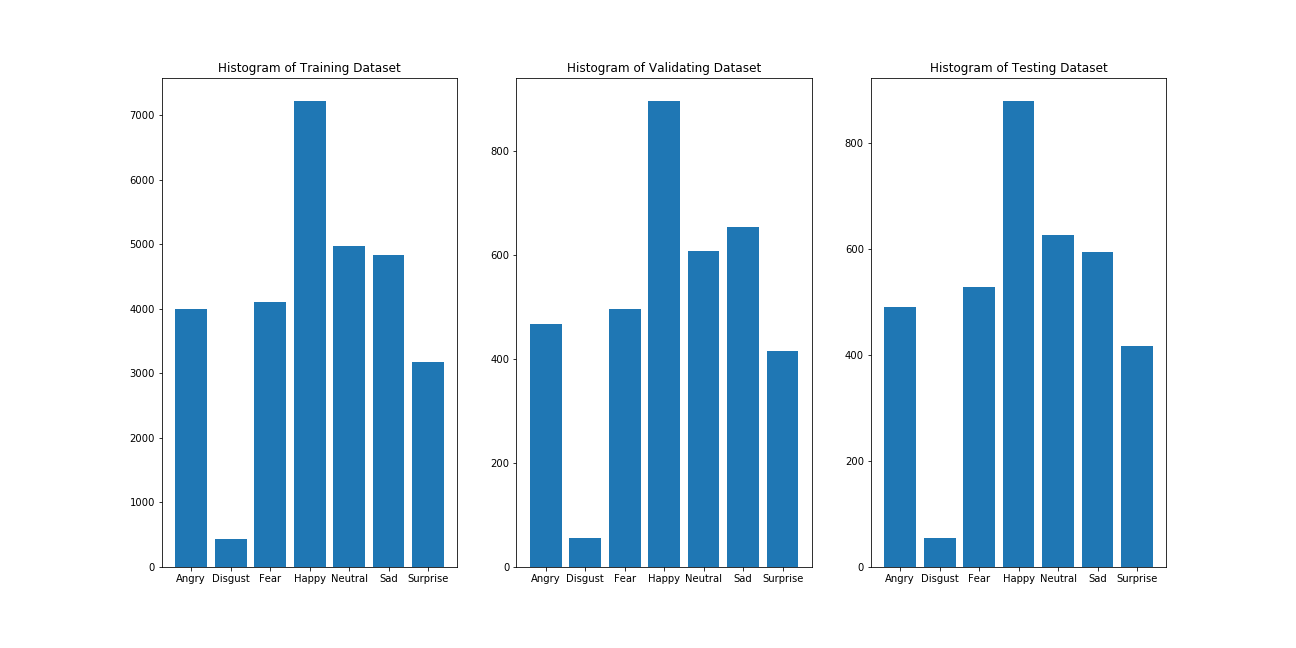
(3) Validating Images:

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/validating_images.png)

(4) Testing Images:

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/testing_images.png)

(5) Data distribution:

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/histogram_images.png)

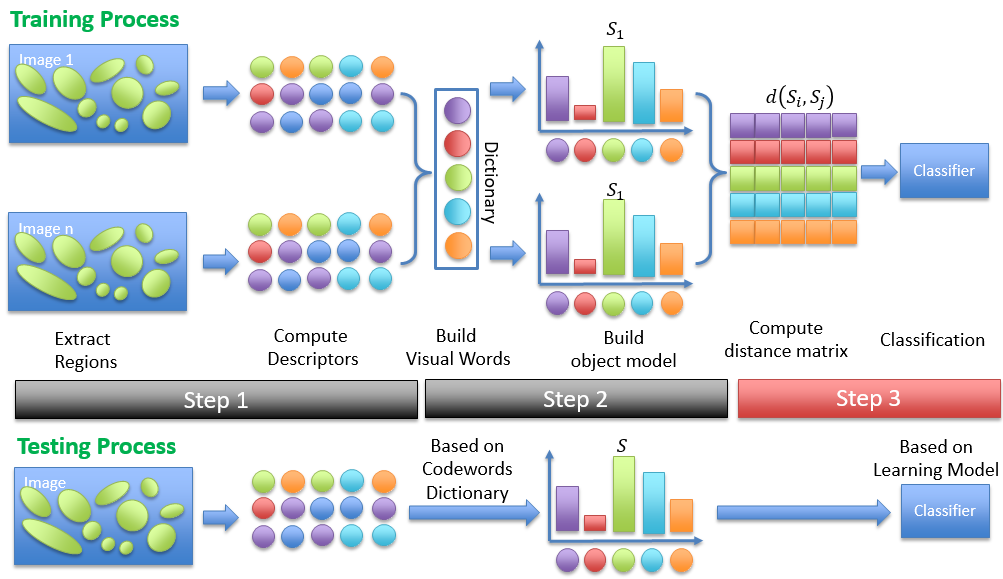
Problem 01: Emotion with Bag-of-Word and Sparse SIFT Feature

(1) Traininig 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

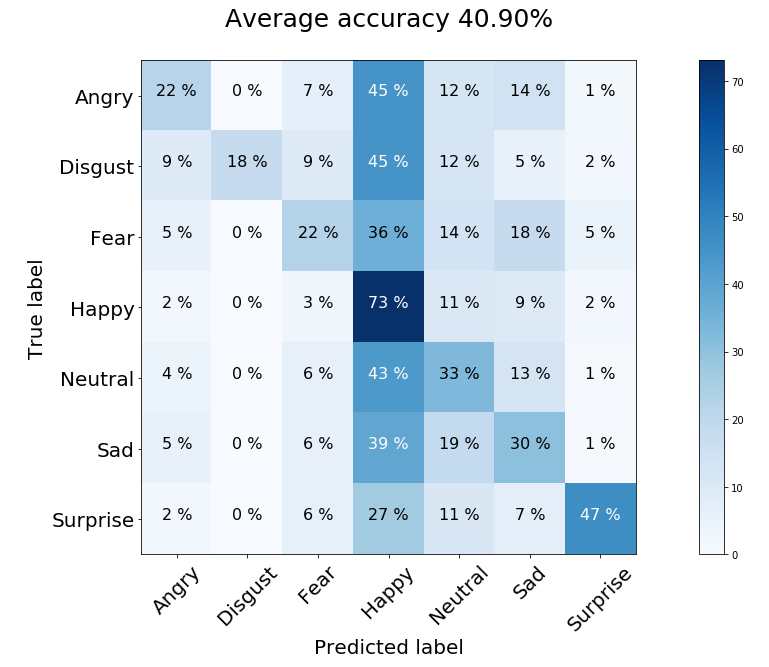
(2) Testing Process:

* 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

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/BagOfWord_Model.png)

(3) Results:

* Number of Codeword Cluster = 4000

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/BagOfWord_Result.png)

Problem 02: Emotion with Deep Learning

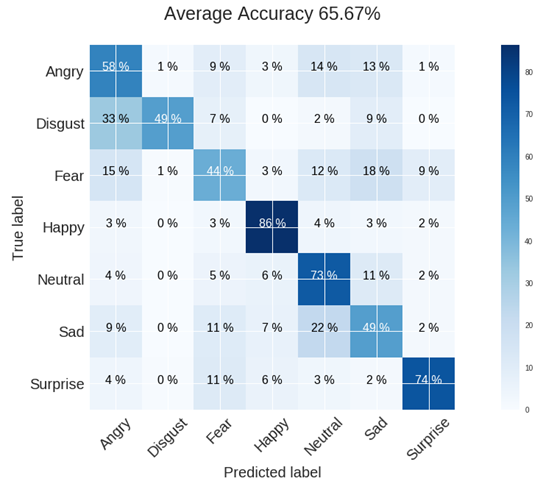
(1) Deep Learning CNN-Like-VGG16 Model:

+ 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 – Flattern, Dense (1024), Dropout(0.5), Dense(7, SoftMax)

(2) Training History:

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/CNN_Training.png)

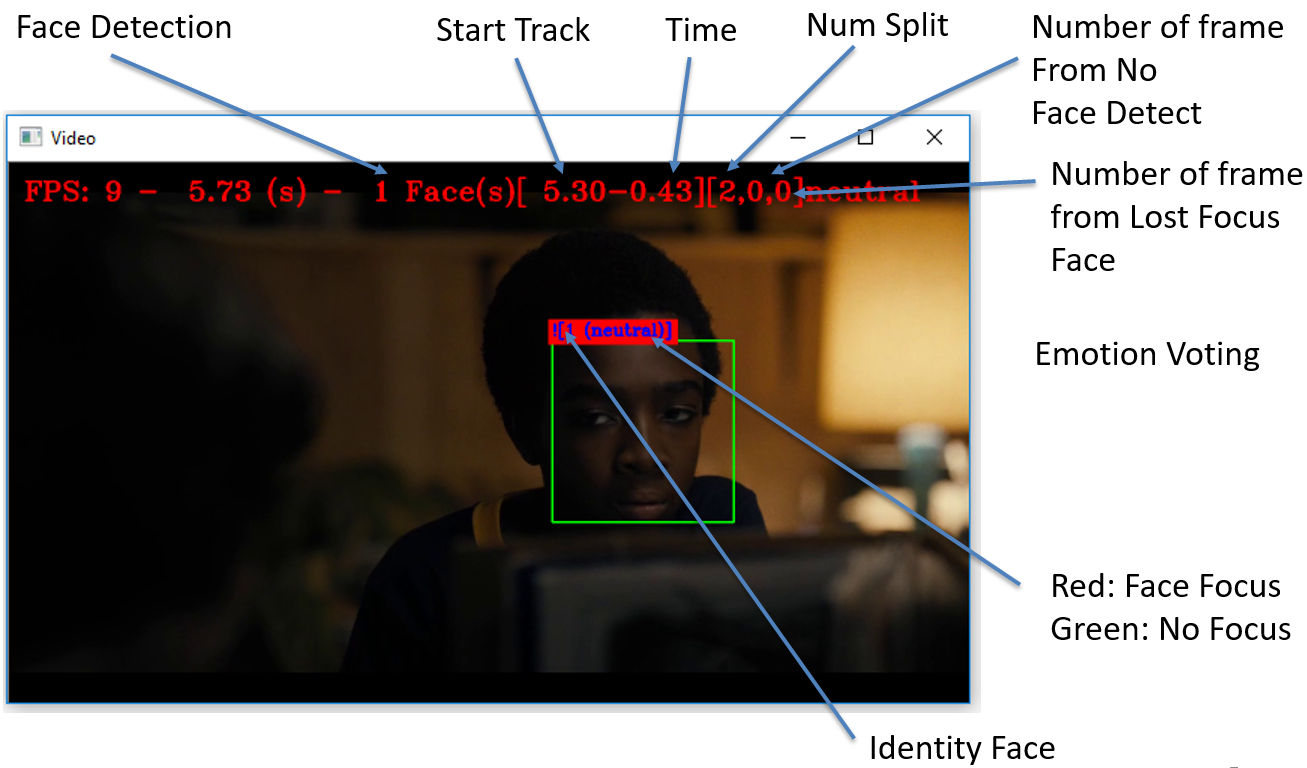
(3) Results:

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/CNN_Result.png)

Problem 03: Video Emotion Extraction

(1) Program Description:

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

[](https://github.com/dntai/dntai_chonnam_computer_vision/blob/master/images/VideoEmotionExtraction.png)

(2) Program Feature:

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

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)

References

Personal information

Do Nhu Tai

Supervisor: Professor Kim Soo-Hyung

Pattern Recognition Lab

Chonnam National University, Korea

E-mail: [donhutai@gmail.com](mailto:donhutai@gmail.com)