ENICARTHAGE
المدرسة الوطنية للمهندسين بقرطاج
Ecole Nationale d'Ingénieurs de Carthage

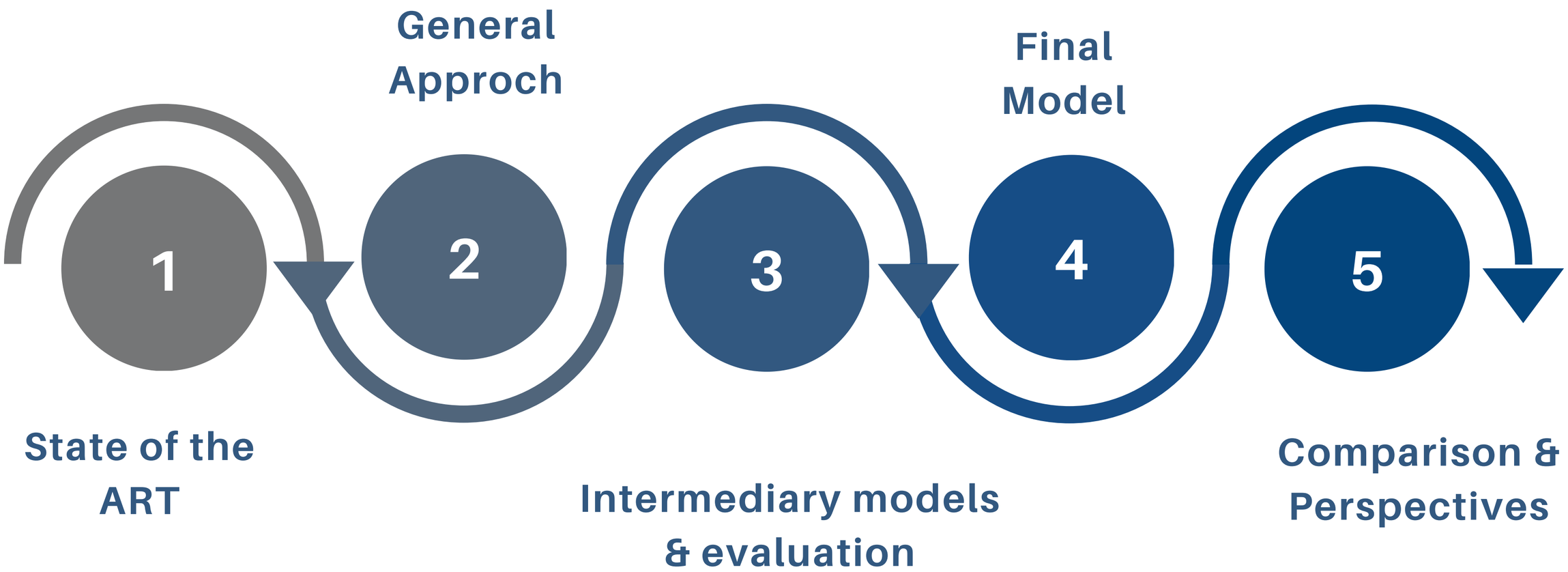
Université de Carthage
جامعة قرطاج

AI Project :

Emotions Recognition

Farah OUESLETI

PLANNING


General
Approch
Final
Model
1
2
3
4
5
State of the
ART
Intermediary models
& evaluation
Comparison &
Perspectives

State of the ART

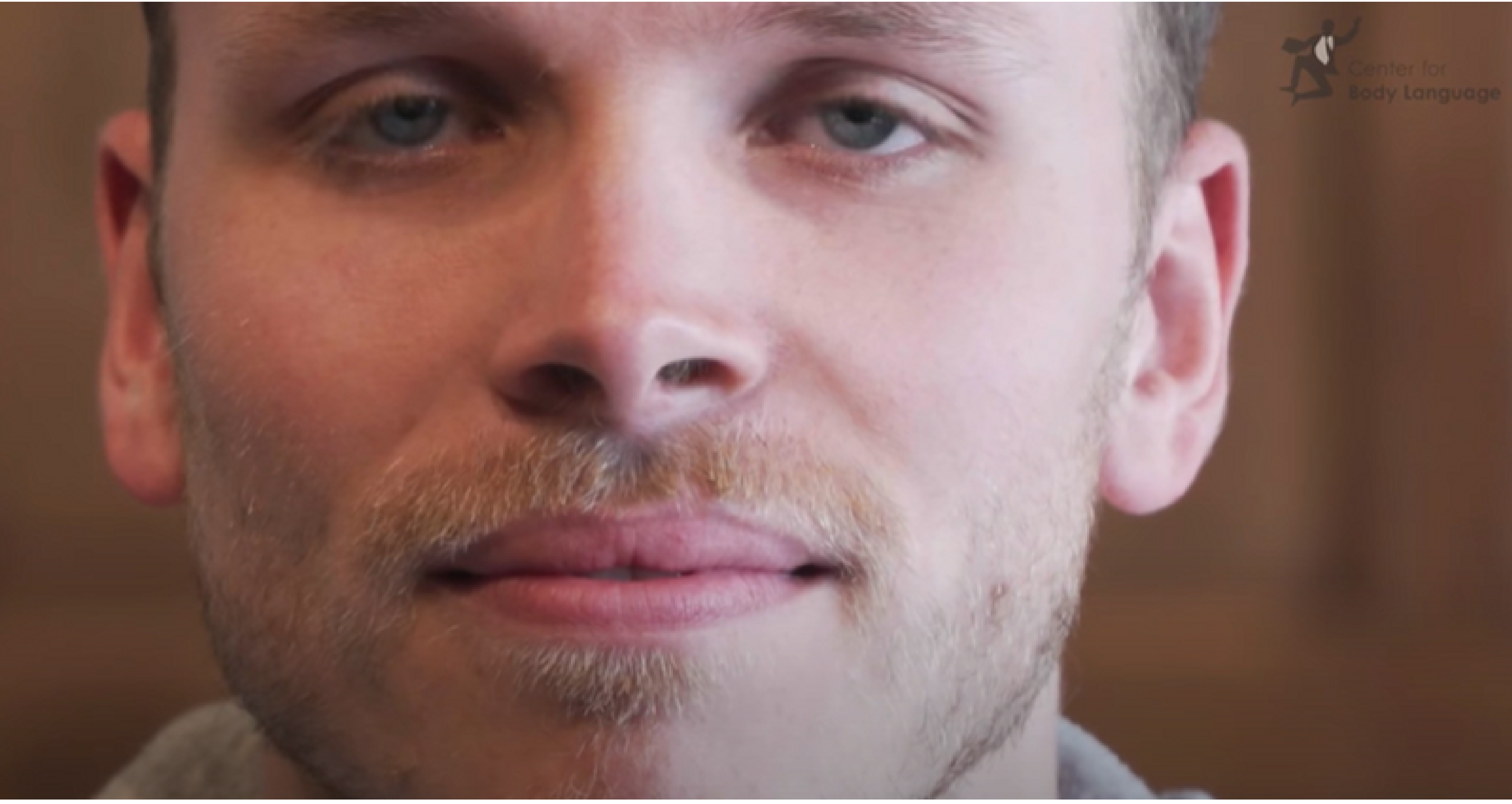
Feature extraction : Relevant

- Mouth ratio to face
- Number of face curves
- Curves around eyes
- Curves around mouth
- Curves around nose
- Upper lip position relative to down lip : size of vertical distance between the two (relative to one lip length)
- Eyebrows pente relative to horizontal axis

Micro expressions :

Facial expressions that last less than 0.5s and are therefore hard to detect.

==> They can serve to distiguish between real emotion and fake or lab made expressions.


Center for
Body Language

Disgust

Sadness

Disgust

Sadness

General Approch

Data augmentation

- **Color shift**
  - **YCbCr**
  - **YES**

--> compare separate training results between

YCbCr

YES

original (greyscale)

Data augmentation

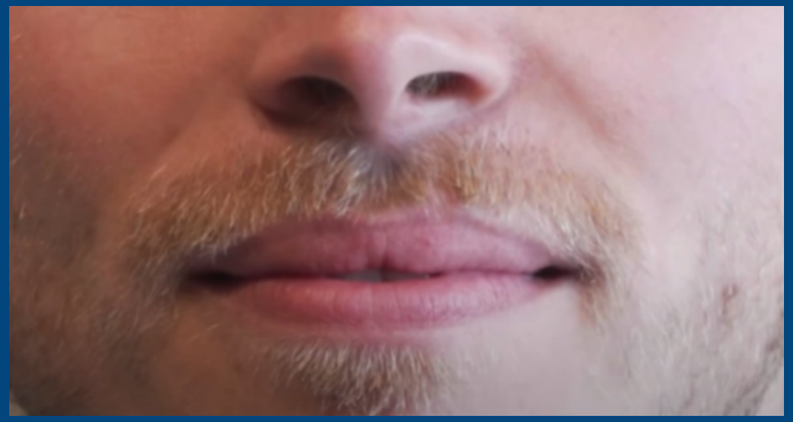
- Gauss Blurring
- Edge Detection
- Data augmentation for ‘Disgust’ emotion

The following choices of techniques adopted in this project (Edge detection filters, noise reducers...) are relative to the Business.

Emotions and facial expressions studies show that :

- Vertical wrinkles around the nose,
- A small upward tilt of the upper lip,
- Slight frowning

==> associated with disgust



- Slight upwaard tilt on the sides of the mouth (wrinkles)
- Small wrinkles on the exterior sides of the eyes

==> indcator of genuine happiness.

- Slight upward tilt on the sides of the mouth
- Various degrees of frowning

==> Associated with sadness.

Gaussian filter

For every pixel, it gives maximum weight to the pixel at hand, and exponentially decaying weights for surrounding values according to how far they are from the central pixel being treated,
which is a very logical way of smoothing out the pictures without creating extremes or associating non representative values to a pixel. (unlike averaging out or taking the maximum value).

Feature Extraction

2 types of methods exist :

- **The appearance feature extraction methods**
- **The geometric feature-based methods**

=> Research has concluded that appearance feature-based techniques achieved better results in terms of accuracy than geometric methods, therfore, we will adopt the first type methods.

Edge detection

These operators have been discarded for the following reasons :

- **Canny :**

it only keeps sharp edges and thins down the result, which won't work well in our case since the pictures are low quality and the wrinkles won't be that visible as opposed to more contrasted pixels (eyes and mouth relative to face will be the most contrasted).

- **Sobel :** Sobel is usually used on RGB colors and performs grayscaling, whereas our dataset is already in grayscale, further grayscaling it will reduce contrast.

=> Chosen operator :

Prewitt

it avoids all the previous inadequecies, and includes a GaussBlurring effect.

Examples of Prewitt filtered

Angry

Disgust

Neutral

Fear

Sad

Happy

Data augmentation

1. We will create an RGB version of the original dataset (the RGB images won't be used for training).
2. The RGB version will be used as an intermediary tool to create the YCbCr dataset and the YES dataset :
   - **YCbCr : a default function is used to go from RGB to YCbCr**
   - **YES : a function is implemented to go from RGB to YES according to transformation matrix.**

3. The test results of the three color spaces will be compared.
4. All the data generated after testing and benchmarking, will be used together in phases for training as an attempt to deliver better test results.

YCbCr

In YCbCr color model, the distribution of the skin areas is consistent across different races in the Cb and Cr color spaces.
As RGB color model is light sensitive so to improve the performance of skin color clustering, YCbCr color model is used.

Its chrominance components are almost independent of luminance and there is non-linear relationship between chrominance (Cb, Cr) and luminance(Y) of the skin color in the high and low luminance region.

Examples of YCbCr

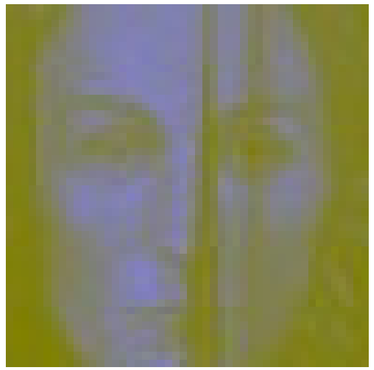
Angry

Happy

Neutral

Fear

Sad



Surprise

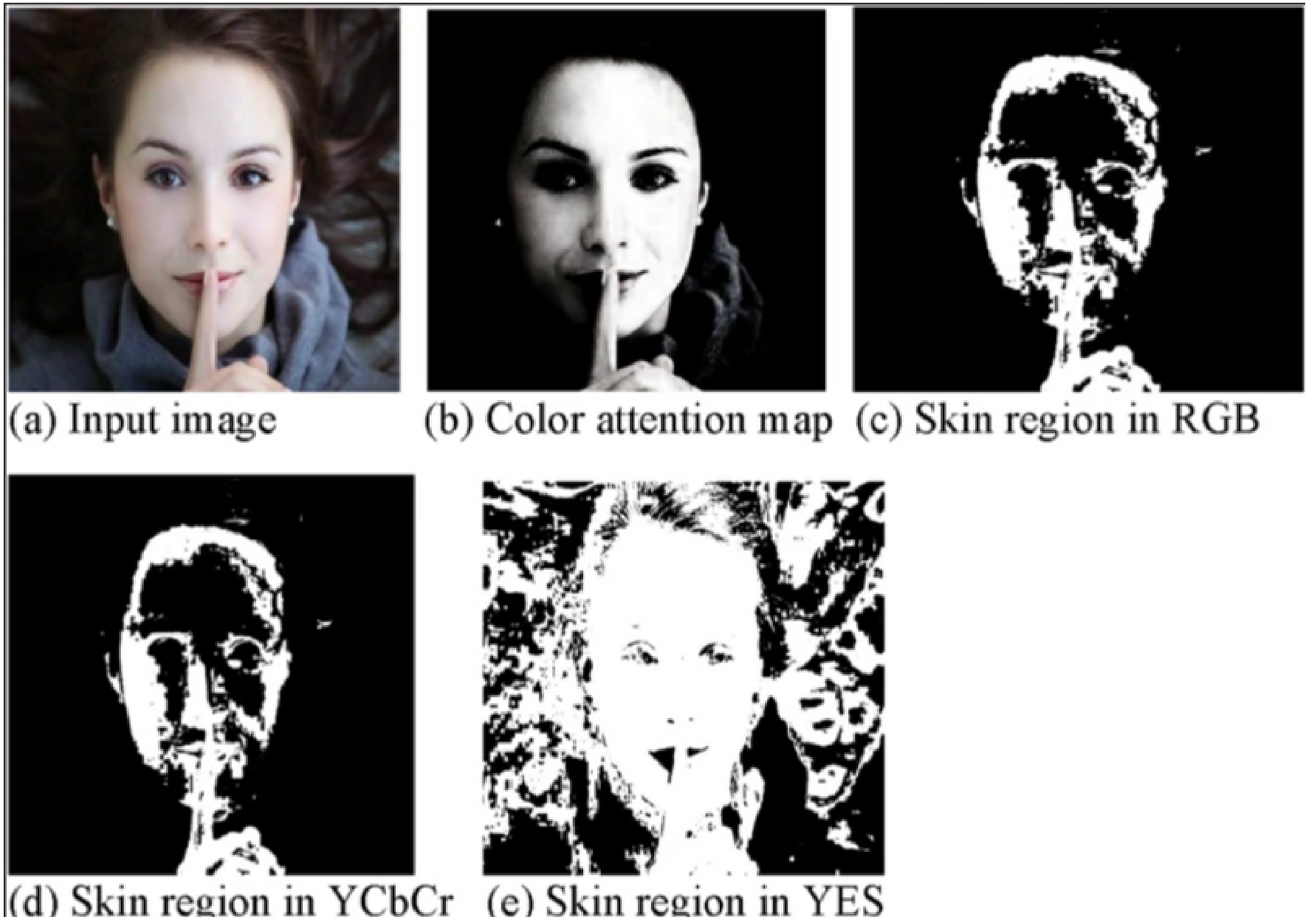
YES

Y represents the luminance channel and E and S denote the chrominance components

It was chosen because :

1. It reduces variations in chrominance due to changes in luminance : in our case, we have a few pictures of dark skinned individuals with very low luminescence, making the face and edges hardly visible for the naked eye, let alone the model.

2. It is computationally efficient the E and S channels can be computed from R, G and B by shifting bits rather than multiplication.



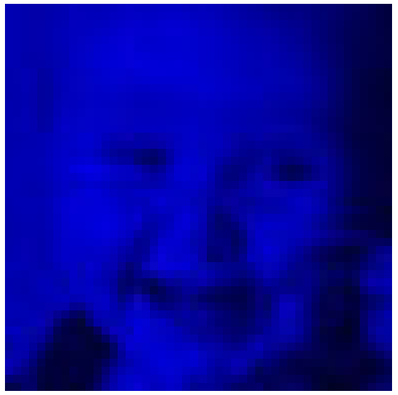
(a) Input image (b) Color attention map (c) Skin region in RGB

(d) Skin region in YCbCr (e) Skin region in YES

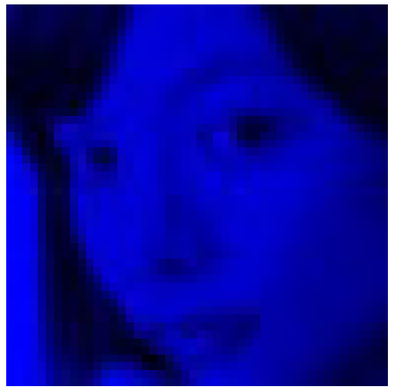
Examples of YES

Angry

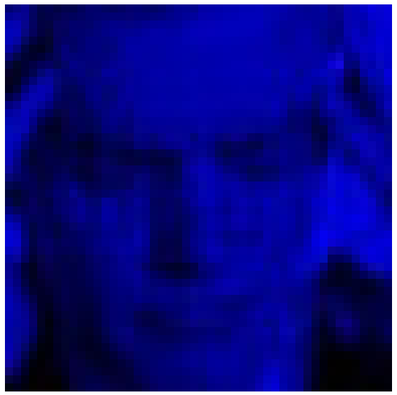
Happy



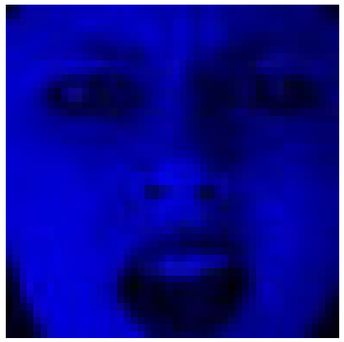
Neutral



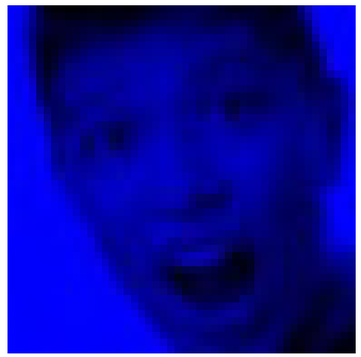
Fear



Disgust



Surprise



Intermediary models & evaluation

A. CNN models

1. Initial CNN model : Overfitting
2. Update 1 CNN model : Very Slow learning
3. Update 2 CNN model : Just Right

Have been tested on GrayScale original data

Testing of Updated model 2 on various datasets :

a. Grayscale Gauss
b. Edge Detection Dataset
c. YCbCr color space images
d. YES color space images

B. Transfer Learning models

1. RESNET model : Overfitting
2. RESNET model : Just right
3. VGG model

Have been tested on GrayScale original data

Final Model

A. Data Intergration : Various datasets integration

B. Final Models

Final models

Final model 1

Concatenation of 2 transfer learning models :

- **VGG**
- **RESNET FINETUNED**

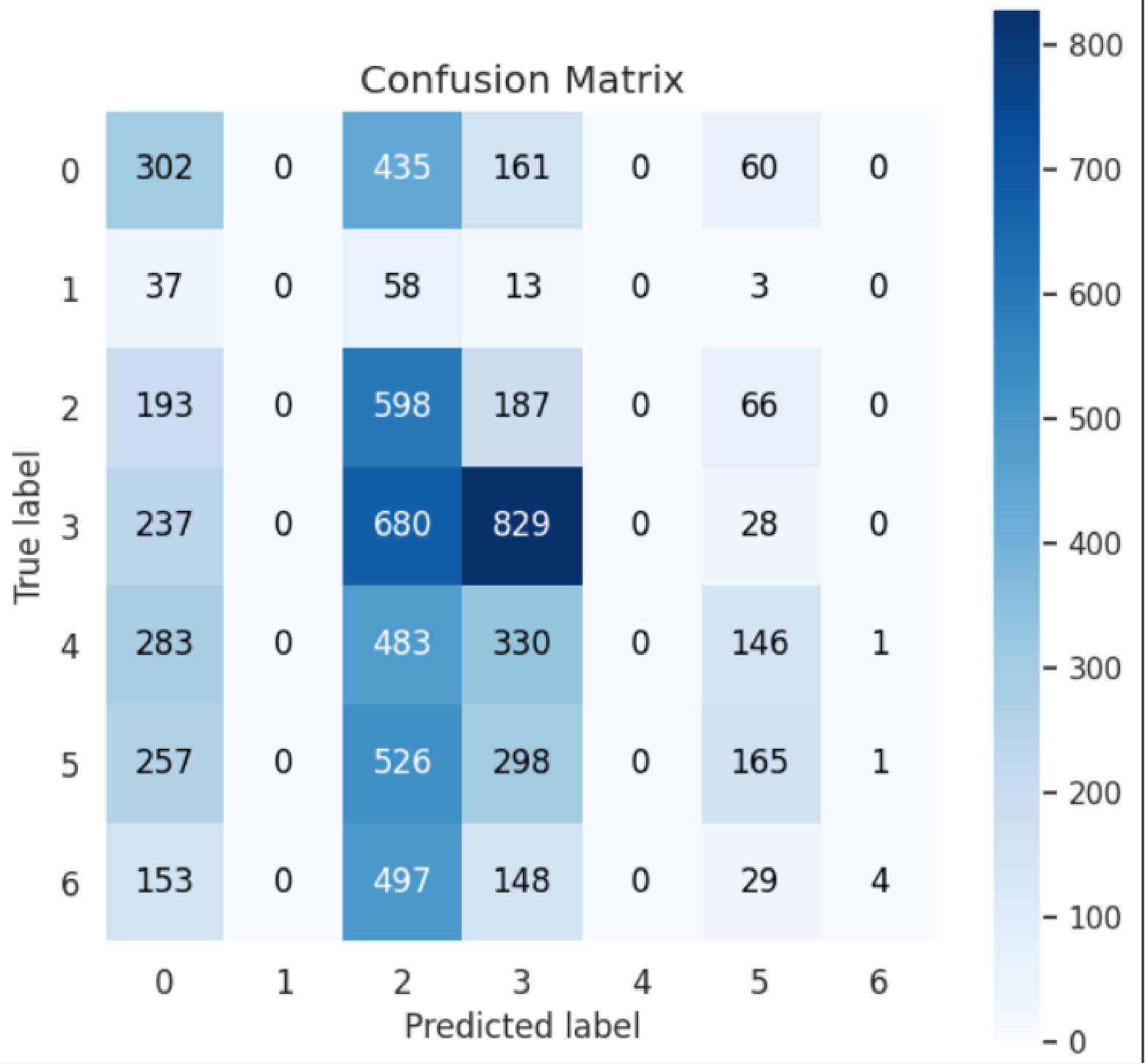
Final model 2

Concatenation of :

- **VGG : Does well in most specifically ‘happy’ and ‘disgust’**
- **CNN updated model 2 : Does well in ‘fear’ detection**

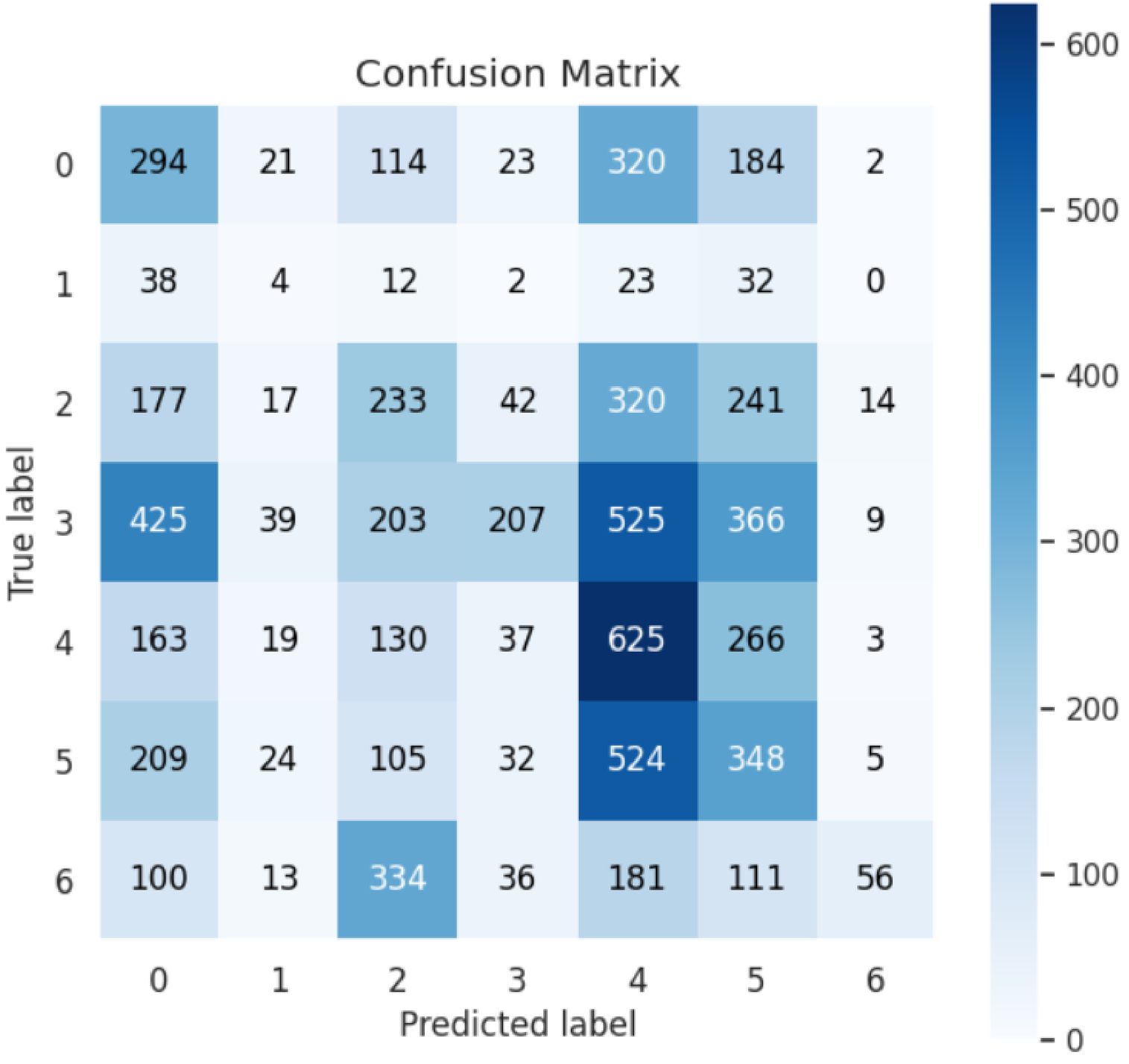
Comparisons & Perspectives

Comparisons



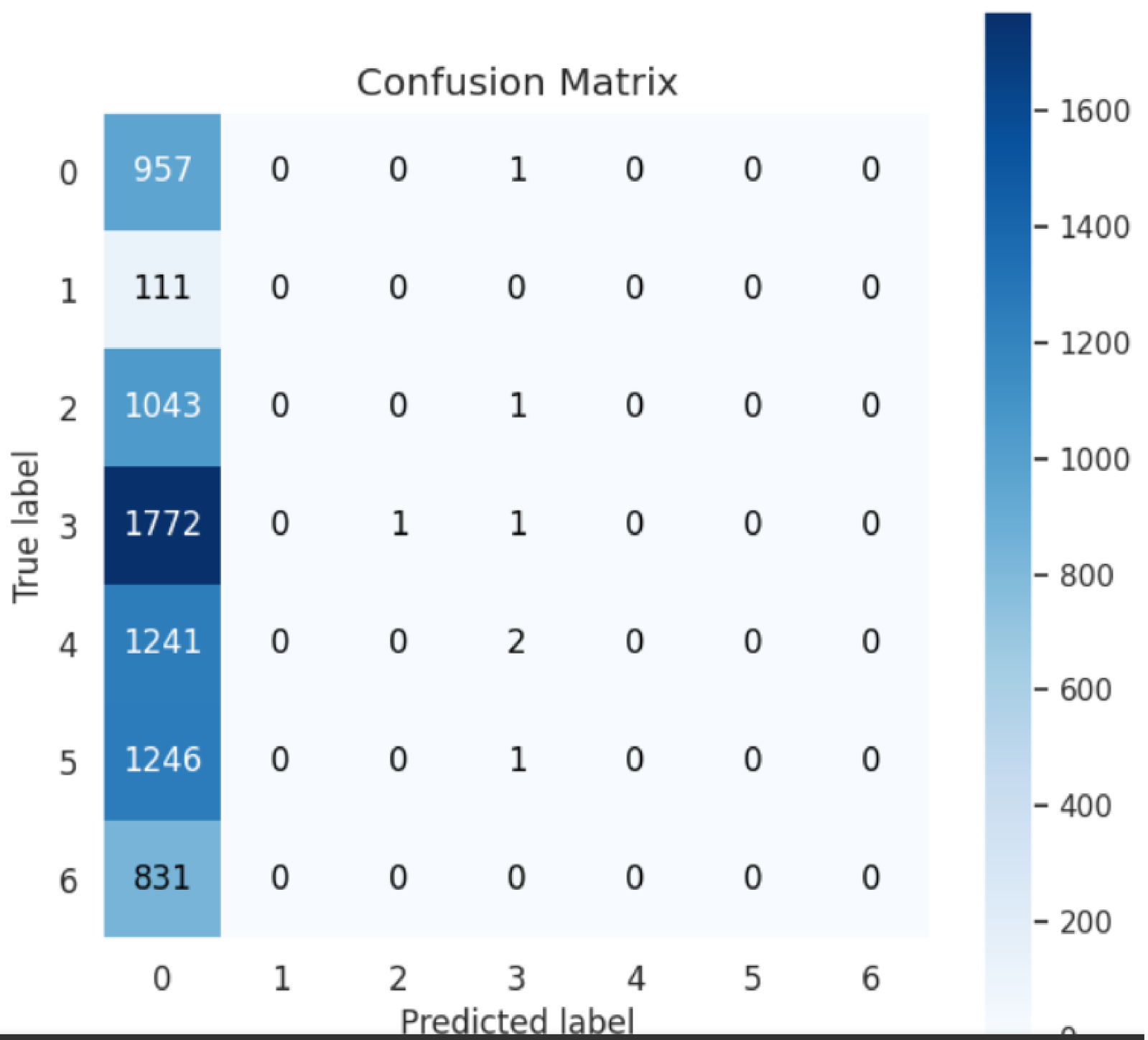
Upadated CNN model on GrayScale

- Gives good results for fear and very good results for happy
- Tends to have high fear recall, many emotions are mistakenly classified as fear.



Upadated CNN model on YCbCr

- Gives good results for neutral
- Tends to mistake some neutral for sad or happy.



Upadated CNN model on YES

- Mistakes Angry for Happy
- Gives good results for angry dataset overall

General Conclusion

- For the same base model, different image treatements with color shift makes te model better at identifying a certain emotion, therefore, including the whole modified dataset as input even to a single model, should give a complete result with good accuracy.
- The final phase is supposed to demonstrate that coupled with other models.

Perspectives

- Using high quality dataset of micro expressions to check model accuracy (Despite being trained on very low quality images)
- Using geometric methods for curves detection
- Using videos of microexpressions as training data
- Using Gabor filters