VIC Project Proposal

Face segmentation

Florian Bettini – Maxime Lutel – Gabriel Muller – Benjamin Pommier

1. Motivation and Problem Definition

Our project aims to implement a multiclass face segmentation algorithm capable of extracting relevant features from a person's figure such as the eyes, mouth, nose, hair and skin through the use of computer vision and machine learning and to apply it to a relevant operation.

1.1. The Problem

In addition to figure's detection, the extraction of face features is a recurring problem in image analysis as it could drastically improve various face-analysis applications by improving the quality of the inputs such as the level of details and information extracted from a single image or videoframe, thus enabling potential new applications.

1.2. Applications

Potential related applications could be head pose estimation, expressions analysis, face matching and identification, gender and age classification as well as deep fake detection. One of those applications will be further develop during the project, in accordance to the features extractions' results.

1.3. Related work

Several papers were written on this subject in order to improve the quality of face recognition and segmentation. The first work on specific features extraction was provided by Yacoob and Davis on hair labelling in 2006 which was later extended to multiple other face features, notably by Khan et al (2015), through the use of random decision forest algorithm and classification. Such features extraction where then uses for multiples face-analysis applications, especially head pose estimation by Benini et al. as well as Khan et al. through the use of random forest classifiers and SVM while Gourier et al. performed face orientation.

2. Methodology

We plan to address the problem in 2 distinct steps:

2.1. Step 1

Use FASSEG dataset (more than 200 labeled faces in multiples poses) to build a model performing two tasks:

- Face segmentation (detect hair, skin, nose, eyes, mouth) as shown in *Figure 1.*
- Head pose estimation (head orientation from -90° to +90°) as shown in *Figure 2.* We will only focus our approach on the yaw rotation, the rotation from left to right.

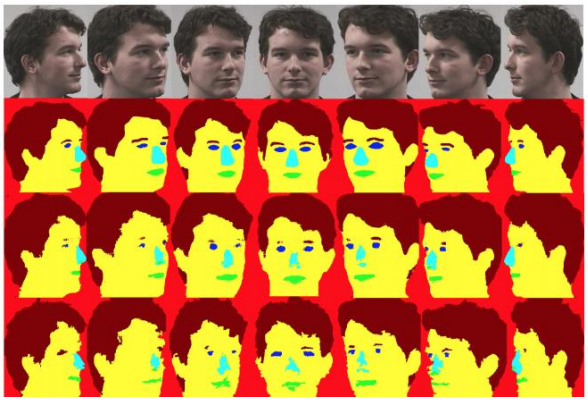

Figure 1. Face segmentation

2.2. Step 2

Use the output of this model to develop a use case among the following:

- Expression analysis
- Gender recognition
- Age prediction
- Face similarities
- Deepfake detection

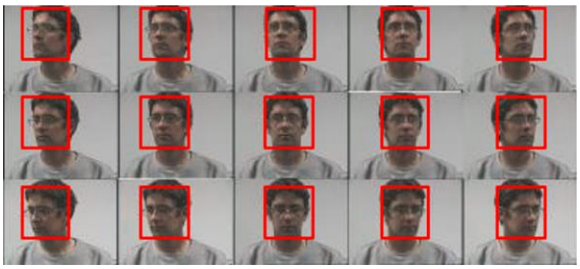

Figure 2. Head pose estimation

Our methodology for step 1 will be based on the one developed by Khalil Khan, Massimo Mauro and Riccardo Leonardi (cf « References » section). For step 2, we plan to extract features from the color maps produced at step 1 as predictors for one of the use cases mentioned above.

3. Evaluation

Evaluating our approach will be based first on its performance (compared to other algorithms) and second on various experiments.

3.1. Performance

The FASSEG dataset (available *here*) will be used for the entire project. As it is already labelled, we will compare the elements we detected with the ground truth labels. 2 methods will be used:

- 1st: as we have less than 200 images, it is possible to compare the different predictions one by one "by hand". So, it will help us visualize absurd or aberrant mouth, head, etc. detection
- 2nd: as it is a classification problem, we will compute a coherent metric to estimate the performance of our model (e.g. Log-loss classification for each pixel, etc.)
- 3rd; as several machine learning models have already been developed (at least for the segmentation part), they will be used as references for our model

3.2. Experiments

Ensuring that our model is accurate enough and scalable is key. Our experiments will be focus on both:

- Optimizing & testing various parameters (as described in the evaluation part of [1])
- Using our algorithms on other unlabeled dataset and analyzing its performance

4. References

Some references were selected in order to support our approach. Mainly from the owners of the FASSEG dataset, we aim to capitalize on their work to develop and implement various head pose detection algorithms.

[1] S. Benini, . K. Khan, R. Leonardi, M. Mauro and P. Migliorati, "Face analysis through semantic face segmentation," *Signal Processing: Image Communication Volume 74*, pp. 21-31, 2019.

[2] K. Khan, M. Mauro, P. Migliorati and R. Leonardi, "Head pose estimation through multi-class face segmentation," *ICIAP*, , vol. LNCS 10485, p. 37–47, 2017.

[3] U. Riaz Muhammad, U. Svanera, R. Leonardi and S. Benini, "Hair detection, segmentation and hairstyle classification in the wild," *Image and Vision Computing* , vol. 71, 2018.

[4] K. Khan, N. Ahmad, K. Ullah and I. Din, "Multicalss segmentation using CRFs," *Turkish Journal of Electrical Engineering & Computer Science*, 2017.

[5] N. Gourier and J. L. Crowley, "Estimating Face orientation from Robust Detection of Salient Facial Structures," *Prima Project*, 2004.

[6] K. Khan, M. Mauro and R. Leonardi, "Multi-class semantic segmentation of faces".

[7] G. Sang, H. Chen and Q. Zhao, "Head Pose Estimation with Improved Random Regression Forests", *Panos Liatsis*, 2015.