

Title of the contribution

September 15, 2015

1 Team details

- Team name
Lab219A
- Team leader name
Bingzhang Hu
- Team leader address, phone number and email
219a Pandon Building
Northumbria University
Newcastle upon Tyne
UK, NE2 1XE
(+44)07597504981
bingzhang.hu@northumbria.ac.uk
- Rest of the team members
Yuming Shen
- Team website URL (if any)
- Affiliation
Dept. Computer Science and Digital Technologies
Northumbria University

2 Contribution details

- Title of the contribution
A deep framework for age estimation
- Final score
Not released yet.
- General method description
STASM+CNN
- References

- Representative image / diagram of the method
- Describe data preprocessing techniques applied (if any)

3 Face Detection Stage

3.1 Features / Data representation

Describe features used or data representation model FOR FACE DETECTION STAGE (if any)

We are using a Haar-Cascade face detector.

3.2 Dimensionality reduction

Dimensionality reduction technique applied FOR FACE DETECTION STAGE (if any)

3.3 Compositional model

Compositional model used, i.e. pictorial structure FOR FACE DETECTION STAGE (if any)

3.4 Learning strategy

Learning strategy applied FOR FACE DETECTION STAGE (if any)

3.5 Other techniques

Other technique/strategy used not included in previous items FOR FACE DETECTION STAGE (if any)

3.6 Method complexity

Method complexity FOR FACE DETECTION STAGE

4 Face Landmarks Detection Stage

4.1 Features / Data representation

Describe features used or data representation model FOR FACE LANDMARKS DETECTION STAGE (if any)

We are using STASM model to detect and represent face landmarks. Only two eye-center landmarks are used in order to crop images and place faces to a fixed position.

4.2 Dimensionality reduction

Dimensionality reduction technique applied FOR FACE LANDMARKS DETECTION STAGE (if any)

4.3 Compositional model

Compositional model used, i.e. pictorial structure FOR FACE LANDMARKS DETECTION STAGE (if any)

4.4 Learning strategy

Learning strategy applied FOR FACE LANDMARKS DETECTION STAGE (if any)

4.5 Other techniques

Other technique/strategy used not included in previous items FOR FACE LANDMARKS DETECTION STAGE (if any)

4.6 Method complexity

Method complexity FOR FACE LANDMARKS DETECTION STAGE

For images from where face and landmark cannot be detected, we just use the whole image as the input of our next stage.

5 Global Method Description

- Total method complexity:
Our framework has low complexity but high efficiency since our network is not very large.
We firstly detect faces in image and then crop and align them into a fixed alignment based on detected landmarks.
We use these cropped faces as the input to our CNN age classifier.
For images from where face and landmark cannot be detected, we manually crop them during training phase, but use the whole image during test phase.
Our CNN classifier gives age scores to each image. We find out the 7 age classes that have the highest scores and make a weighted average as the final age output.
- Which pre-trained or external methods have been used (for any stage, if any)
We use a pre-trained CNN network focusing on faces so called Age/Gender.net based on Adience OUI dataset.

- Which additional data has been used in addition to the provided ChaLearn training and validation data (at any stage, if any)
We use Adience OUI dataset as our external training source.
- Qualitative advantages of the proposed solution
We reach a score of 0.47 on the validation set provided.
We also evaluate our frame work on Morph dataset as an alternative, which achieves 3.8 years in MAE. This outperforms the state-of-the-art.
- Results of the comparison to other approaches (if any)
CCA with BIF features has been explored on the validation set, reaching about 0.65 in score and 7.2 years in MAE.
- Novelty degree of the solution and if is has been previously published
We combine some published methods together, and make some modifications so that the frame work can work correctly.

6 Other details

- Language and implementation details (including platform, memory, parallelization requirements)
For face and landmark detection, we programme in C++.
We use Caffe to build our CNN network.
We use the Matlab interface of Caffe to evaluate our performance.
- Human effort required for implementation, training and validation?
Yes.
- Training/testing expended time?
Fine-tuning on the pre-trained CNN model costs 5 to 10 hours in our experiments.
We spent about one month on this challenge in total.
- General comments and impressions of the challenge? what do you expect from a new challenge in face and looking at people analysis?
We haven't yet explore the influence of deviation in the training set and hopefully we can go deeper into it in the future.