Apparent Age Estimation using LGBP features and Cascade Ridge Regression

September 11, 2015

1 Team details

- **Team Name:** Notts CVLab
- **Team Leader:** Mercedes Torres Torres
- **Address:** Office B86, School of Computer Science, Jubilee Campus, University of Nottingham, NG8 1BB, United Kingdom
- **Phone number:** +44 7778 498988
- **E-mail:** mercedes.torrestorres@nottingham.ac.uk
- **Team Members:** Michel Valstar, Brais Martinez, Yorgos Tzimiropoulos
- **Affiliation:** Computer Vision Lab, University of Nottingham

2 Contribution details

- **Title of the contribution:** Apparent Age Estimation using LGBP features and Cascade Ridge Regression
- **Final score:** Mean epsilon of 0.46 over validation set.
- **General method description:** A four-stage age estimation system based on LGBP-feature extraction and uses a 3-level cascade to estimate the ages of unseen images.
- **References:** To be found at the end of the document.
- **Representative image:** A diagram of our method is shown in Figure 1.

3 Face Detection Stage

Our first stage is the Face Detection Stage. The purpose of this stage is to locate a face within an image and to encase it in a bounding box.

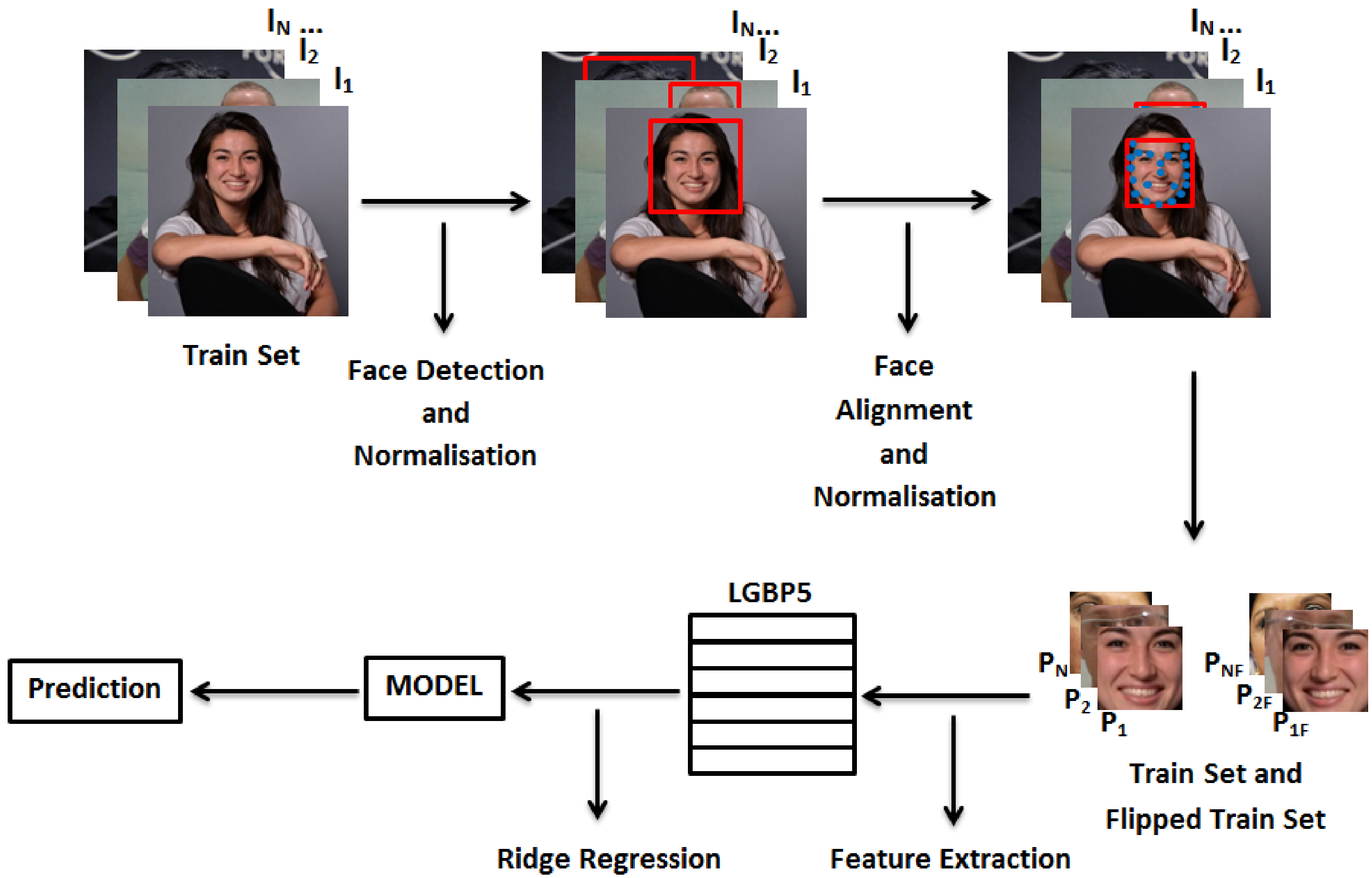

Figure 1: Diagram of our method

3.1 Features / Data representation

We have followed the approach proposed by Zhu & Ramanan in [6], which entails using Histogram of Oriented Gradient features [1] for face detection.

3.2 Compositional model

The compositional model follows a Tree-based deformable parts model, as proposed in by Zhu & Ramanan in [6] and further developed by Asthana et al in [1].

3.3 Learning strategy

Our learning strategy follows the method explained in the Zhu & Ramanan [6] paper. However, our face-detection model has been re-trained with in-the-wild images [1].

3.4 Method complexity

The authors did not provide the method complexity.

4 Face Landmarks Detection Stage

After the faces have been localised and normalised using Procrustes analysis, the next step is to localise landmarks within the faces [4].

4.1 Features / Data representation

A parametric global shape model and a parametric part-based appearance model are used as data representation schemes [4].

4.2 Dimensionality reduction

PCA is applied as a dimensionality reduction technique [4].

4.3 Learning strategy

The key idea is to compute an averaged Jacobian $\hat{J}(k)$ from a set of training examples from which the facial appearance variation is projected-out. The averaged projected-out Jacobian, denoted as $\hat{J}_P(k)$, is then used to compute an averaged projected-out Hessian and descent directions [4].

4.4 Method complexity

The complexity of this stage is $O(nN)$ [4].

5 Feature Extraction Stage

After the faces have been aligned and normalised once again with Procrustes, we extract LGBP [3] features.

5.1 Features / Data representation

Local Gabor Binary Patterns (LGBP) features [3] are extracted by first creating a set of Gabor magnitude response images (one for every filter in a filter bank) and then applying an LBP operator to each of them. This has been shown to be very robust to illumination changes and misalignment. Additionally, in order to increase the number of samples to be used during training, we extract LGBP features from both the images in the original train set and the images obtained from flipping the train set.

5.2 Dimensionality reduction

As in previous steps, once the features have been extracted, we apply PCA to reduce their dimensionality.

5.3 Method complexity

LGBP feature extraction has a complexity of $O(n^2)$.

6 Learning Regressor Model Stage

Once features have been extracted, we train our regressor following a cascaded ridge regression method.

6.1 Features / Data representation

We use the result of applying PCA over the training set with the original images and the flipped images.

6.2 Learning strategy

We apply cascaded Ridge Regression [2] to train our regressor. Our cascade has three levels. To train the regressor of the first level, the complete train set is considered. However, for the second and third level of the cascade, a regressor is trained for each possible age. Each of these regressors is trained with a subset of the original training samples according to an interval, which ranges between the age in question minus the standard deviation of the predictions of the samples with that age and the age in question plust the standard deviation of the predictions of the samples with that age.

6.3 Method complexity

Ridge Regression has a complexity of $O(n^3)$.

7 Global Method Description

- **Total method complexity:** $O(n^3)$
- **Which pre-trained or external methods have been used (for any stage, if any):** No pre-trained methods or models have been used for this challenge. External code that implements the approach that we chose for Face Detection [1], is used in the fist stage. This code was made available by the zetahu & Ramanan in [5] . All the other code has been developed by members of the team.
- **Which additional data has been used in addition to the provided ChaLearn training and validation data:** None.
- **Qualitative advantages of the proposed solution:** Our system makes use of LGPB5 and cascaded regression.
- **Results of the comparison to other approaches (if any):** No comparison to other approaches has been made.
- **Novelty degree of the solution and if is has been previously published:** The main novelty stems from using cascaded regression with subsets of training samples. This subset is calculated according to the predictions of the previous levels of the cascade.

8 Other details

- **Language and implementation details**: Our system has been implemented using MATLAB R014b. We have used PC equipped with a Intel Core i3 Processor, 8GB of RAM and Windows 7.
- **Human effort required for implementation, training and validation?** Two people were involved in the implementation, training and validation.
- **Training/testing expended time?** Training the cascaded regressor a Train Set of 2,415 images (plus another 2,415 flipped images) takes 9:36 hours. The testing time for a Test Set of 1,079 images, is of 2 hours.
- **General comments and impressions of the challenge? what do you expect from a new challenge in face and looking at people analysis?** In general, I found the challenge quite interesting.

References

[1] Akshay Asthana, Stefanos Zafeiriou, Shiyang Cheng, and Maja Pantic. Robust discriminative response map fitting with constrained local models. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3444–3451. IEEE, 2013.

[2] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

[3] Thibaud Senechal, Vincent Rapp, Hanan Salam, Renaud Seguier, Kevin Bailly, and Lionel Prevost. Combining aam coefficients with lgbp histograms in the multi-kernel svm framework to detect facial action units. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 860–865. IEEE, 2011.

[4] Georgios Tzimiropoulos. Project-out cascaded regression with an application to face alignment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3659–3667, 2015.

[5] X. Zhu. Face Detection, Pose Estimation and Landmark Localization in the Wild. `http://www.ics.uci.edu/~xzhu/face/`, 2008. [Online; accessed 2015-06-30].

[6] Xiangxin Zhu and Deva Ramanan. Face detection, pose estimation, and landmark localization in the wild. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2879–2886. IEEE, 2012.