Correlation between Professional Domain and Facial Features based on Face Clustering

Bithiah Yuan

October 16, 2019

1 Introduction

A significant source of information and attributes can be derived from the human face by non-verbal communication [10]. As a result, facial features have been studied extensively in the social-science domain to predict success in reaching reputable leadership positions. In particular, studies have shown that certain facial features contribute to higher salaries and more prestigious employments for CEOs. In application, the relationship between facial characteristics and social attributes can provide an more powerful objective indicator for organizations to idenity and select effective leaders within their domain than broad facial cues such as attractiveness and competence. Results have shown that a human judge can identify business, military, and sports leaders from their faces with above-chance accuracy. However, these results are biased and do not imply the actual leadership qualities of a person [13].

Caused by behaviour experiments from human judgement, the the research of the social attributes and facial features in the social-sciences are limited in scalability, consistency, and generalization. For example, prior familiarity to the faces of the study and personal preferences can affect the results. Therefore, a growing number of social trait judgment studies have been extended and refined to computer vision and machine learning research due to the capability of using massive datasets and large-scale processing capacity [10].

Through a computational framework, [10] examined the relationship between facial traits and the social construction of leadership by a trained model that can predict the outcomes of political elections based on the perceived social attributes of a person's appearance. The results indicate that similar methods can be used to predict behavior in a broad range of human social relations, such as mate selection, job placement,and political and commercial negotiations [10].

Clustering analysis is an unsupervised learning technique that groups data points into clusters based on their similarities. It is useful in grouping a collection of unlabeled data with similar nature into clusters. [19] investigated clustering a large number of unlabeled face images into individual identities present in the data [19]. The workflow shown in Figure1 consists of obtaining face representations of a collection of unlabeled date by a deep neural network. The choice of clustering algorithm then groups the face images according to their identity.

Motivated by the researches in computational social trait judgment and [19], the following paper aims to examine the correlation between a person's profession based
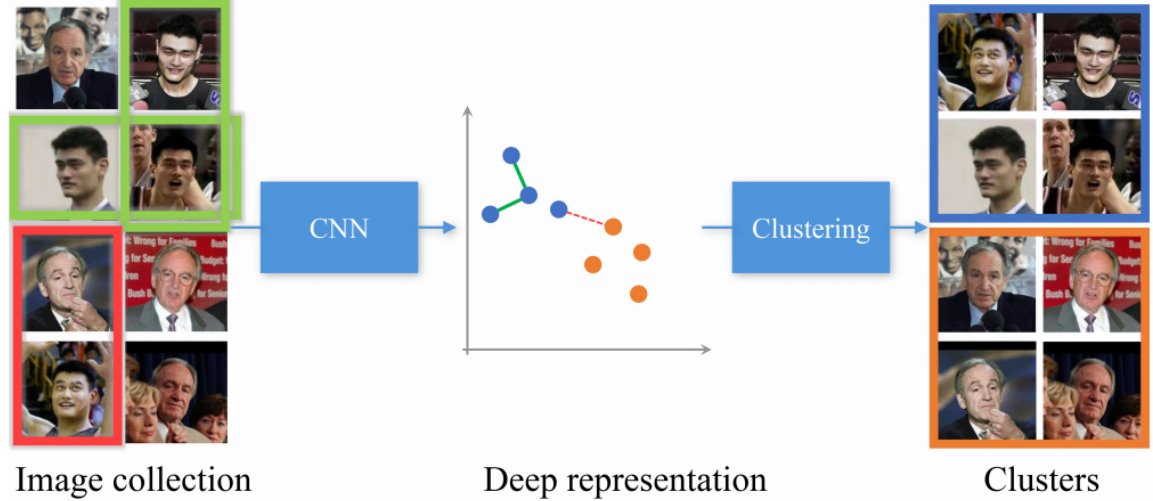
Figure 1: Face clustering workflow [19]

on their facial features through clustering face images. The clustering problem consists of the face representation and similarity metric of the face images and the choice of clustering algorithm [19]. Due to the importance of the underlying face representation in face clustering, this paper further compares different open-source state-of-the-art feature extraction methods based on deep learning.

2 Background

2.1 Face Recognition

Face recognition focuses on identifying or verifying the identity of subjects in images or videos [16]. Challenges occur in real world (in-the-wild) images due to variations ranging from head poses and illumination conditions to aging and facial expressions. Recently, deep learning methods based on convolutional neural networks (CNNs) were able to achieve very high accuracy by learning robust features due to the availability of large-scale faces in-the-wild datasets on the web [16].

Face recognition systems are usually composed of the following 4 steps [16]:

1. Face Detection:
   Detect the position of the faces in an image and returns the coordinates of a bounding box for each face as shown in Figure 2.

2. Face Alignment:
   Find a set of facial landmarks with the best affine transformation that fits a set of reference points located at fixed locations in the image as shown in Figure 3.

3. Face Representation:
   Trasnform the pixel values of a face image into a discriminative feature vector.

4. Face Matching:
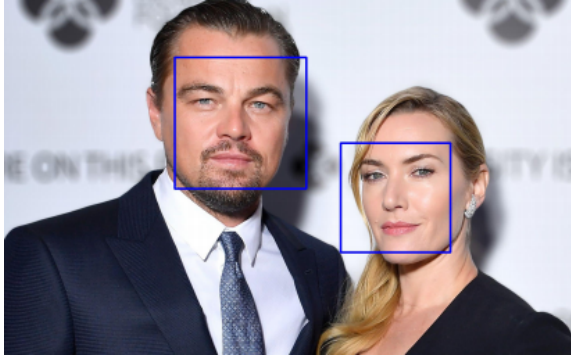   Compute similarity scores from feature vectors.
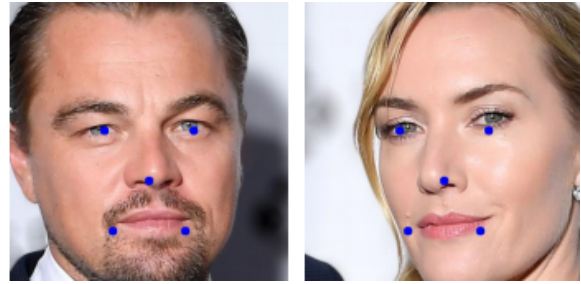
2

Figure 2: Face Detection [16]

Figure 3: Face Alignment [16]

2.2 Feature Extraction

2.3 Clustering

3 Experiment

3.1 Dataset

3.2 Evaluation

4 Conclusion

References

[1] *2.3. Clustering.* https://scikit-learn.org/stable/modules/clustering.html. 2011.

[2] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. *OpenFace.* http://cmusatyalab.github.io/openface/. 2016.

[3] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. *OpenFace: A general-purpose face recognition library with mobile applications.* Tech. rep. CMU-CS-16-118, CMU School of Computer Science, 2016.

[4] David Arthur and Sergei Vassilvitskii. "K-Means++: The Advantages of Careful Seeding". In: vol. 8. Jan. 2007, pp. 1027–1035. DOI: 10.1145/1283383.1283494.

[5] Adulwit Chinapas et al. "Personal Verification System Using ID Card and Face Photo". In: *International Journal of Machine Learning and Computing* 9 (Aug. 2019), pp. 407–412. DOI: 10.18178/ijmlc.2019.9.4.818.

[6] Navneet Dalal and Bill Triggs. "Histograms of Oriented Gradients for Human Detection". In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005)* 2 (June 2005).

[7] Jiankang Deng, Jia Guo, and Stefanos Zafeiriou. "ArcFace: Additive Angular Margin Loss for Deep Face Recognition". In: (Jan. 2018).

[8] Adam Geitgey. *Face Recognition.* https://github.com/ageitgey/face_recognition. 2017.

[9]   Jia Guo and Jiankang Deng. *InsightFace: 2D and 3D Face Analysis Project.* `https://https://github.com/deepinsight/insightface`. 2019.

[10] J. Joo, F. F. Steen, and S. Zhu. "Automated Facial Trait Judgment and Election Outcome Prediction: Social Dimensions of Face". In: *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015, pp. 3712–3720. DOI: `10.1109/ICCV.2015.423`.

[11] Davis E. King. "Dlib-ml: A Machine Learning Toolkit". In: (July 2009).

[12] Davis E. King. *High Quality Face Recognition with Deep Metric Learning.* `http://blog.dlib.net/2017/02/high-quality-face-recognition-with-deep.html`. 2017.

[13] Christopher Olivola, Dawn Eubanks, and Jeffrey Lovelace. "The many (distinctive) faces of leadership: Inferring leadership domain from facial appearance". In: *The Leadership Quarterly* 25 (Oct. 2014). DOI: `10.1016/j.leaqua.2014.06.002`.

[14] Charles Otto, Dayong Wang, and Anil Jain. "Clustering Millions of Faces by Identity". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PP (Apr. 2016). DOI: `10.1109/TPAMI.2017.2679100`.

[15] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[16] Daniel Saez Trigueros, Li Meng, and Margaret Hartnett. "Face Recognition: From Traditional to Deep Learning Methods". In: (Oct. 2018).

[17] David Sandberg. *facenet.* `https://github.com/davidsandberg/facenet`. 2018.

[18] Florian Schroff, Dmitry Kalenichenko, and James Philbin. "FaceNet: A unified embedding for face recognition and clustering". In: June 2015, pp. 815–823. DOI: `10.1109/CVPR.2015.7298682`.

[19] Yichun Shi, Charles Otto, and Anil Jain. "Face Clustering: Representation and Pairwise Constraints". In: *IEEE Transactions on Information Forensics and Security* PP (June 2017). DOI: `10.1109/TIFS.2018.2796999`.

[20] Kaipeng Zhang et al. "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks". In: *IEEE Signal Processing Letters* 23 (Apr. 2016). DOI: `10.1109/LSP.2016.2603342`.