Face Recognition: Challenges and Issues in Smart City/Environments

Praveen G.B
PathPartner Technology Pvt Ltd
Bangalore, India
praveen.gb@pathpartnertech.com

Jayachandra Dakala
PathPartner Technology Pvt Ltd
Bangalore, India
jayachandra.d@pathpartnertech.com

***Abstract*—Smart city challenges, like increased traffic, risk to public safety, effective law enforcement and the smart environment challenges on improving personalized services such as health care and home environment need person identification. Face recognition has proved to be useful and amicable bio-metric for smart city and smart environment challenges. In this paper, we review the use cases pertaining to smart city and smart environment leading to use case specific requirements on Face recognition. We describe the open challenges in anti-spoofing and standardization to make face recognition a fool proof system. For applications with demand for low power, we show that with proper considerations on training data, cost function, and model architecture we could build a low-complex CNN model with reasonable accuracy (91.4% on LFW data set).**

***Index Terms*—Face recognition, smart city, smart environment, smart interaction**

I. Introduction

Rapid urbanization brought lot of challenges [1] to public life like increased traffic, risk to public safety and effective law enforcement. We can call them smart city challenges. At the same time with intelligent computing becoming ubiquitous [2], every device that humans interact with becoming smarter day by day. This brings up the questions on how to personalize services like health care, home environment, etc. In both the smart city challenges and smart environment or smart interaction challenges, recognizing people is a must.

Of the wide range of bio-metrics like finger print recognition, iris recognition, face recognition, etc., face recognition (FR) is favored the most for the reasons like: a) FR is the natural way of recognizing people, b) is non-invasive, and c) low cost. For the same reasons, demand for FR has grown rapidly in both smart cities and smart environments use cases. In fact, a recent market study [3] predicted that FR system will have a market worth of $7.0 billion by 2024.

In this paper, we review the use cases and attempt to summarize the FR requirements and open problems. Particularly, in Section II, we review the use cases and categorize the requirements. In Section III, we present our approach on how to trade-off complexity and accuracy of a FR system. In Sections IV and V, we talk about the open problems in face anti-spoofing and FR system standardization, respectively. Finally, in Section VI, we summarize the paper with future directions.

II. Use cases and Requirements

In this section, first we briefly describe different use cases that would need face recognition, then we attempt to categorize the use cases into different classes based on the type of requirements they would impose on face recognition.

A. FR use cases

We categorize the use cases into four broad classes and list some of them under each class. Note that this is not an exhaustive review, but acts as a reference to understand the different requirements on face recognition. We have also tagged use cases proposed by us.

1) Health Care: Hospitals and elderly care homes are investing heavily on face recognition to monitor sick patients. Some of the aspects covered are as follows.

- Patient check-in process
- Patient tracking
- Managing medicine distribution
- Diagnosis. Ex: Face2gene [4] uses face recognition to diagnose certain disease
- Staff identification
- Retrieve critical information of patients in emergency. Ex: In case of accidents, help public (or) ambulance drivers to identify the blood group of victim and their details [proposed]

2) Smart Homes and Smart Rooms: Use face recognition to ensure maximum security and to customize user specific services. Examples:

- Smart doorbell with face recognition- to identify any entrants
- Customization in travel industry as well as smart rooms using face recognition and identifying unique users to personalize surroundings to them
- Person-specific recommendation systems and services for smart homes
- Automatic floor selection in Elevators for aged and disabled people like blind [proposed]

3) Security, Access Control & Law Enforcement:

- Easy boarding and security checks at airports
- Jaywalking violations, as used in China [5]
- To book traffic violations by rental transport users (such as bounce, vogo, drivezy) [proposed]

978-1-7281-3187-0/20/$31.00 ©2020 IEEE

- Identifying victims of human trafficking (E.g. to identify missing children)[proposed]
- Criminal identification in the context of environment [proposed]
- Identification of men in female-reserved coaches (or) women-only areas like hostels [proposed]
- To build a unified penalty system (FR-based fine collection system for all kind of violations : traffic police, public transport usage without valid ticket) [proposed]

4) Marketing & Retail:

- Payments on public transports (E.g. metro trains, public buses, toll gates, electricity bill payments)
- Face recognition for retail stores to make shopping more convenient
- Face recognition for supermarkets to track shoppers with suspicious behavior and identify shoplifters
- Identify users based on face recognition for customized advertisements and promotions

B. Requirements and Challenges

In general, any FR system should be able to generate unique feature representation (acting as an ID) for a given person such that it is invariant to intra-person variations (illumination, pose, expression and occlusions) and quite different from all inter-class (among different persons) variations. Further, use case dependent requirements are as listed below.

1) RGB Vs NIR: Illumination variation is a key challenge in vision-based applications. Applications with low-illumination or poorly illuminated environments (such as driver monitoring systems, night-time surveillance applications) need Near Infra-Red (NIR) based input. For practical use cases, NIR-based FR model is not a completely solved problem due to a) lack of sufficient training data and b) lack of face texture variations in NIR space compared to RGB input.

2) Large-scale Vs Small-scale: Depending on the number of people in a given use case, FR system can be classified as either large-scale or small-scale. Applications such as personalization and authentication systems would fall under small-scale, whereas, FR systems for law enforcement agencies and surveillance systems would fall under large-scale in which a test image needs to be compared against too many people. Developing a model that can generate unique feature vector against both intra class and inter class variations for large-scale deployment and yet keeping the complexity within practical limits is a challenge. As an example, in Section III, we show how to trade-off complexity and accuracy of an FR model suitable for small-scale usage.

3) Images Vs Video: Applications such as photo retrieval systems, payment kiosks, access control systems operate on images. In such cases, we will only have one test image to be either verified against or identified from existing data base. This demands a FR system with better precision. On the other hand, surveillance systems and smart home based applications make use of video based input. In this case we will have more than one image of a person to authenticate. Hence even with a reasonably accurate FR system, post-processing of results across multiple frames can result in higher accuracy. Additional methods such as multi-face trackers can reduce computational complexity and help to track people in real-time.

4) Computational complexity: Model complexity selection is completely driven by use case. Small-scale use cases and deployment on power-efficient devices such as body cam, mobile phones need low complex models. For applications such as post-analysis and for deployment on servers, high-complex models can be used. Model selection with right trade-off between complexity and accuracy is a key challenge as running time can vary on different platforms.

III. Face Recognition

Face recognition can be solved using either CNN based or non-CNN based approaches. To handle variations such as illumination, pose & occlusion, non-CNN approaches may need additional methods [1] and are not recommended for usage in applications at scale.

To build a feature representation for FR, we need training dataset with enough variations, CNN architecture complex enough to capture variations needed and appropriate loss function. Existing approaches uses too-deep networks trained with softmax loss or shallow networks with locally connected(LC) layers both of which have higher complexity. Due to usage of unshared weights, LC layers results in higher model-size and higher memory bandwidth. It also limits the usage with deeper networks as the complexity may increase multi-fold. Hence LC-layer based CNNs cannot be used in low-latency applications. For coarse-grained classification tasks, softmax loss is good enough to separate the classes in feature space. However, for fine-grained classification tasks like FR, either fine-tuning with triplet loss [6] or joint-training of center loss [7] with softmax loss has proven to result in better accuracy. In use cases with higher number of faces to be processed, low-complex model is a key requirement. Hence, we propose PP-FRNet constituted with 7 convolution layers and one fully-connected layer.

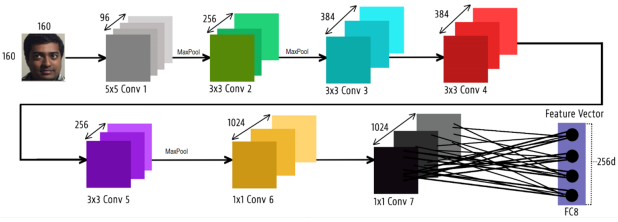


Fig. 1. CNN Architecture of proposed PP-FRNet

We compare our proposed CNN w.r.to InceptionResNet-v1 in terms of computational complexity and accuracy in Fig-2. We use Labelled Faces in Wild (LFW) test data set for comparison, which has 5749 people with total of 13,233 images. Total of 6000 pairs are formed and used as test set. Computational complexity is measured on nVidia TitanX Pascal.

In video-based input, we have used feature representation

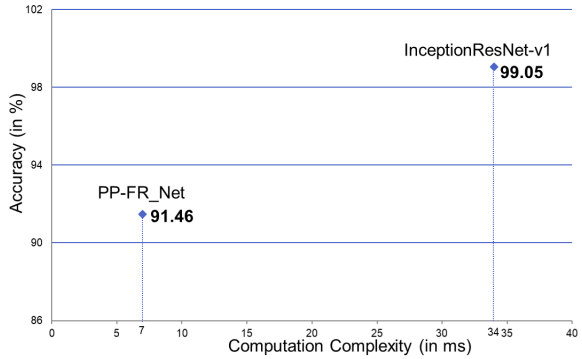


Fig. 2. Comparison of FR model trained with InceptionResNet-v1 vs PP-FRNet

across multiple frames for recognition and further used face tracker which resulted in reduced complexity and overall improvement of system accuracy. Tracker is reset with recognizer results based on two conditions a) when face is not available and/or b) major change in scale & size of the tracked face.

IV. Face anti-spoofing

Along with other challenges to be addressed in FR systems, spoof attack is a major challenge when deployed for practical use. Print attack, eye-cut photo attack, warped photo attack, replay attack, 3D mask attack [8] are few of the familiar spoof attacks used to degrade FR performance in access control systems and application that demands zero false recognition rate. Wearable face projectors, face illuminator, face paints, facial accessories, clothing/head scarf are some of the recent trends in spoof attacks targeting to avoid being recognized, and hence, reducing the recognition rate. Instead of relying on a single FR model, an additional model to detect spoof and assist in decision making would result in better accuracy.

V. Need for standardization

Can we rely completely on FR systems? Can they be misused? As there is no standardization for FR systems, there are privacy and ethical usage concerns like making the data public, misusing the FR system to target specific set of people based on ethnicity or gender. Citing the same concerns, FR usage has been banned in California (refer to bill AB1215 [9]) and other cities. To address such concerns a standardization working group, titled IEEE P7013 [10], has been formed. P7013 is majorly focusing on standardizing associated metrics and concerns related to bias and inaccuracy. Based on our understanding, we make the following recommendation under three broad categories for standardizing FR systems, which we believe would make it complete and comprehensive.

1) Ethical usage:

- In critical applications like law enforcement agencies, FR should be used as assistive technology and furthermore with intervention of humans
- Using confidence score along with recognition results in decision making

2) Privacy:

- Strict laws against personal data misuse and recommendation to store metadata than using video input or images.

3) Bias & Decision making:

- Defining exhaustive set of use cases and corresponding variations (or) challenges for each use case
- Defining relevant metrics and its acceptable range for each use case
- Test data set exhaustively representing use case. Dataset should also have uniform representation on ethnicity, gender along with other variations listed above. Based on use cases, test set also represent data with spoof and look-alike to evaluate the robustness of the method.
- Test datasets should not be collected from publicly available database or from internet, as it may overfit if already used in training

VI. Future Directions

In this review, we looked at different use cases and their requirements on FR system. Then, we briefly discussed the challenges in anti-spoofing and the need for standardization. We have also proposed a low-complex model that can be used in small-scale real-time applications.

In summary, to develop and deploy FR systems for a wide range use cases, following would be the broad list of open problems to be addressed:

- Development of NIR based scalable and robust FR models
- Development of FR models that are scalable yet low-complex for large-scale deployment
- Development of FR systems that are robust against variety of fake face attacks
- Effective and comprehensive standardization of FR systems addressing ethical, privacy, and bias and decision making issues

References

[1] Q. Chen et al., "A Survey on an Emerging Area: Deep Learning for Smart City Data," in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 3, no. 5, pp. 392-410, Oct. 2019.
[2] A. Pentland and T. Choudhury, "Face recognition for smart environments," in Computer, vol. 33, no. 2, pp. 50-55, Feb. 2000.
[3] https://www.marketsandmarkets.com/PressReleases/facial-recognition.asp.
[4] https://www.face2gene.com/
[5] Tao, Li. 2018. "Jaywalkers under Surveillance in Shenzhen Soon to Be Punished via Text Messages." South China Morning Post. March 27,2018. https://www.scmp.com/tech/chinatech/article/2138960/jaywalkers-under-surveillance-shenzhen-soon-bepunished-text
[6] arXiv:1503.03832
[7] Wen Y., Zhang K., Li Z., Qiao Y. "A Discriminative Feature Learning Approach for Deep Face Recognition." In ECCV 2016.
[8] arXiv:1904.02860
[9] https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill textunderscore id=201920200AB1215
[10] https://standards.ieee.org/project/7013.html