PARTIAL FACE RECOGNITION USING DYNAMIC FEATURE MATCHING AND CNN

Mr.R.Raj Bharath, M.E. (Ph.D.)
Associate Professor,
Computer Science and Engineering,
Manakula Vinayagar Institute of Technology,
Puducherry, India.

K. Arunkumar
Computer Science and Engineering,
Manakula Vinayagar Institute of Technology,
Puducherry, India.

S. Diwakar
Computer Science and Engineering,
Manakula Vinayagar Institute of Technology,
Puducherry, India.

P. Dineshkumar
Computer Science and Engineering,
Manakula Vinayagar Institute of Technology,
Puducherry, India.

Abstract— As the computer generation grows, security and safety are being an active topic. The most active and most advanced problem is face detection and face recognition. Nowadays, detection and recognition of frontal face images are achieved in all computer fields, from identification to authorization and authentication. Still, the problem is to deal with a partial face. Sometimes the obtained picture may be occlusions, large-viewing angles, or out-of-view. Identifying a partial face and recognition isn't an easy task. We propose a technique in this paper to detect the partial face and recognize them using various steps. Our method helps to identify the partially appearing faces with very low computational complexity. We couple the detected facial segment, which allows detecting partially visible faces. Once a partial face is detected, the facial sections can be used in many applications. The partially detected probe is sent to a face recognition model to discover the person. We use FCNs (Fully Convolutional Networks) and SRC (Sparse Representation Classification) combined using DFM (Dynamic Feature Matching) to handle the recognition of partial face regardless of different sizes. (*Abstract*)

Keywords—Partial Face Detection, Partial Face Recognition, Deep learning, Fully Convolutional Network.

I. INTRODUCTION

With new technological advancements, images have become various convenient and better tools of expression. Millions of photos stored in different cloud storage and it is used across social networks every day. Retrieving and organizing these photos impacts user experience and is very challenging. Images are two-dimensional projections of three-dimensional things like faces. The face is not a unique determined object. There are more than millions of different faces, and each of them can assume a variety of deformations. Inter-personal changes can be due to identity, race, genetics, or ethnicity, whereas intra-personal modifications can be due to distortions, look, facial hair, aging, etc. We deal with the difficulty of multi-view face detection in a given image. Face identification has been an ongoing study is for the past two decades. Hardy discovered techniques are available which enable detection of upright faces in real-time with very moderate computational complexity. These involve cascade simple to complex face classifier. Several modifications of these have been implemented in digital cameras and smartphones.

Figure 1. Partial faces are obtained from different environments. The face may be located partly out of cameras field of view or captured in different postures or occluded by mask, scarf.

It is challenging to detect and recognize partially occluded face images. Usual mask identification or scarf detection algorithms deal with mixed algorithms like feature-based algorithms [1], [7], [9] and training-based algorithms. Purposes based on facial characteristics exploit the information of facial features such as eye, mouth, etc. to choose whether there is no occlusion on the probe or not. The partial face detection [3], [4] deals with the quality of image, conditions on differing lighting, the partial occlusion of the image, and atypical face detector would consequently be able to identify the appearance of any face under any set of light conditions, upon any environment. The detected partial face is further passed to face recognition model to recognize the person.

Multi-scale Region-based CNN [2], [4], [5] is used to deal with partial face recognition. At first, it decomposes an incomplete face probe into several region proposals while obtaining features of each region proposal by deep convolutional neural networks. The partial face is recognized by using region-to-region matching. Although MR-CNN [3] delivers extraordinary performance on some partial face images, it needs the presence of certain facial parts and pre-alignment. Sliding Window Matching (SWM) [6] introduces an alternate solution for partial face recognition through installing up a sliding window with a similar size as the test image, which is used to search for the most relevant region within all gallery images. However, the sliding process is performed at the image-level, which is computationally expensive. In this research, we propose a partial face recognition system: Dynamic Feature Matching (DFM) [16], [6], which could manage partial faces of inconsistent sizes without additional pre-processing, including computational efficiency and high accuracy. Fully Convolutional Networks [10] is suitable for input images with irregular dimensions while making spatial feature maps with similar sizes of the input pictures. At first, Fully Convolutional Networks [4], [5], [7] is used to obtain spatial feature maps of given probe faces and gallery. The final pooling layer is practiced as a feature extractor despite the scale/size of the input face. After that, followed by Sliding Window Matching, we fixed a sliding window of the equivalent size as the examination feature maps to decay the gallery feature maps inside some gallery sub-feature maps in feature-level. All gallery sub-feature maps and the dimension of probe feature maps are identical. Provided a probe, we disintegrate the whole gallery feature maps into sub-feature maps matching to the size measurement of the probe feature maps externally computing gallery feature maps frequently. This method increases the rate of speed over 10× compared to Sliding Window Matching [17]. Finally, Sparse Representation Classification (SRC) [16] is applied to obtain alignment-free matching. Sparse Representation Classification uses a one-sample per class approach, delivering partial face-matching with alignment-free matching.

The sparse solution provides a feasibility design where the probe feature plans are linearly described by these gallery sub-feature maps. SRC reduces the restoration error without requiring a constraint on the choice of gallery sub-feature maps. Therefore, some unusual sub-feature maps could be likely chosen to provide the least rebuilding objective error. To fix this issue, we combine sub-feature maps selection constraints to SRC. Consequently, related sub-feature maps bring much more in-depth study in feature development, and the possibility of mismatching is significantly decreased. The effectiveness of this approach is verified in our analyses.

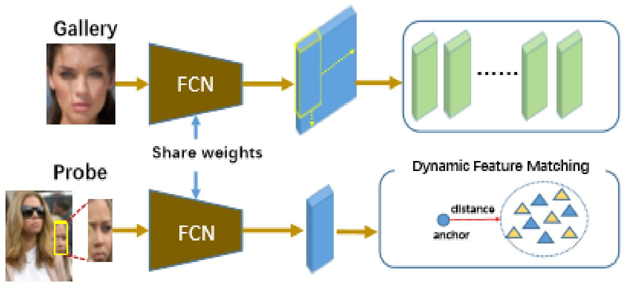

Figure 2. The framework of our proposed partial face identification: DFM (Dynamic Feature Matching). A fully Convolutional Network is executed based on the VGGFace model.

The partial face recognition part is divided into three subdivisions

• Dynamic Feature Matching (DFM) [16], [2], [3], [5], merges Fully Connected Networks with SRC, delivering maximum recognition accuracy, and performance in computational efficiency.

• The suggested approach betters the present partial face recognition algorithms on some face databases. Also, DFM obtains competitive performance on partial person re-identification, and it can be extended to other computer vision obstacles.

• DFM [16] can not only run for holistic face pictures still additionally can deal with partial faces of inconsistent size with alignment-free.

In the following sections of this paper, we explained detail about our related work, proposed work, result, analysis and our conclusion.

II. Related Work

Face detection is one of the studies in computer vision. In the early period (before 2000), many practical performances and studies of the face detection were not enough until the proposed work of Viola and Jones, and these two are the first people to draw a rectangular-shaped box for the face. But, this method had many disadvantages more than the advantages because of its large size. At present, face recognition is used for security purposes such as identification and authentication. This is a feature used to replaces the use of passwords, pin numbers and also fingerprint. Because people facing some problems like digit limitation in a pin number, even a password has some limitations and even fingerprint scanner too has problem such as wet fingers, dry skin and also it has no way when the person gets wounded in his/her fingers. In order overcome and improve the security authentication system face recognition comes into the role. While we use for identity or authentication, face recognition feature differs a lot from usual face recognition systems.

The 24 x 24 image has totally 160,000 number of haar like specification and also and it is not capable of handling After finding faces such as wild and frontal to overcome the above defect many people have made lots and lots of work and introduced much more complicated specifications such as SURF, HOG, SIFT and ACF [8], [9], [11]. After that, they introduced another new feature called NPD which has the ability to differentiate the intensity of pixels between two pixels. One of the popular method is Dlib which supports classifier in the face detection system. Another model which are greatly discussed is enlarging the robustness of detection. Another simplest method is combining multiple detections models which must be trained in different views each separately. It is called as an applied multiple deformable models that captures faces in different angular views. These models require testing and training where it consumes lots of time and it provides less performance is less. In 2002 a person named Garcia et al. had introduced a model developed with help of neural network to capture and identify the semi -frontal human faces in the critical pictures. In 2005 a person called Osadchy et al had trained the models with convolutional neural networks [19], [20] for the detection of face. Lots of face detection and recognition work has been carried out since it doesn't require any kind of human effort for the detection and recognize faces as many technical people have involved in the invention of many face detection and recognition methods.

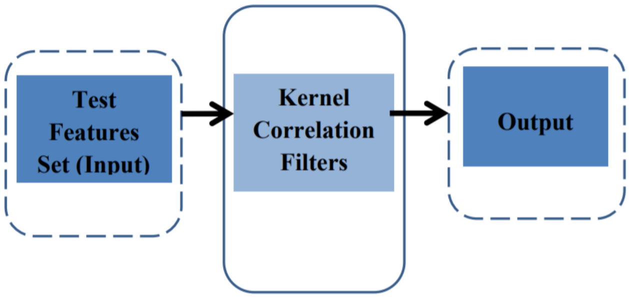

Fig. 3.0

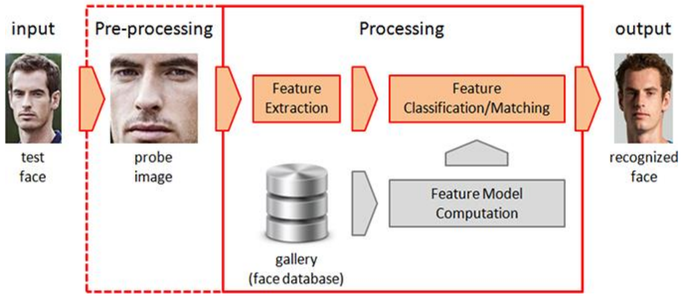

Fig. 3.1

Figure 3.0 & 3.1. Represent the existing system architecture.

III. OUR APPROACH

We divided our approach into two halves, they are

1. Partial face detection
2. Recognition of partially detected probe.

1. *Partial Face Detection*

For partial face detection, Alexnet fine-tuned at first. We use the AFLW dataset contains 24k face annotation and 21k images. The number of positive samples was improved by randomly examining sub-windows of the pictures. The Intersection of Union(IOU) with ground truth was observed with all such images. Images having higher than 60% IOU were used as positive samples. This ended in 200k positive samples and 20 million negative samples. Whole photos were resized to 227 × 227. For the fine-tuning of Alexnet, a group size of 128 copies containing 34 positive samples and 94 negative examples was used for 50k iterations. Either a sliding window approach or a region-based approach can be followed to detect partial faces.

In this paper, we describe the selection of a sliding window approach because of its less complexity.

Database	# landmarked imgs	# landmarks	# subjects	image size	image color
Caltech 10,000 Web Faces	10,524	-	-	-	color
CMU/VASC Frontal	734	6	-	-	grayscale
CMU/VASC Profile	590	6 to 9	-	-	grayscale
IMM	240	58	40	640x480	color/grayscale
MUG	401	80	26	896x896	color
AR Purdue	508	22	116	768x576	color
BioID	1,521	20	23	384x286	grayscale
XM2VTS	2,360	68	295	720x576	color
BUHMAP-DB	2,880	52	4	640x480	color
MUCT	3,755	76	276	480x640	color
PUT	9,971	30	100	2048x1536	color
AFLW	25,993	21	-	-	color

Figure 4. The difference between other face dataset and AFLW

The method is not arranged and discusses essential face detection problems by a fast filtering strategy. We create a broad Multi Key-point Description (MKD) [7], [9], recognize multiple key Scale Invariant Feature Transformation (SIFT) [2], [5] marks in all images, and measure the proper descriptors to provide partial face recognition. The MKD-SRC algorithm profitably implemented the database of gallery class information and suppressed fraudulent similarity, seeking the most economical representation of all images in the gallery.

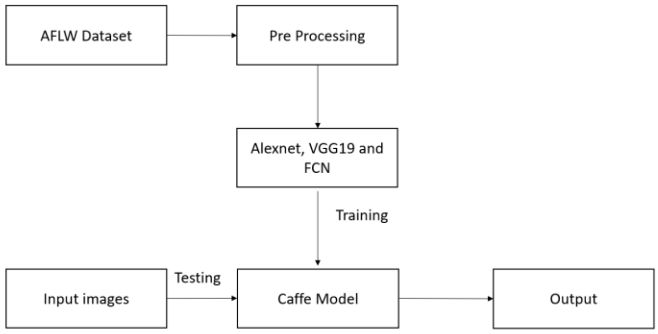

Figure 5. The Architecture of proposed partial face detection

This face detector contains three fully connected layers and five convolutional layers. We prefer to use python with the Caffe framework.

Our method has Alexnet, VGG19 and FCN of layers as below

Convolutional Layer 1(Activation - Rectified Linear Unit; Filter - 18×18×1×60; Bias - 60; Stride - 1×1)

Max-pool Layer 1(Filter - 2×2; Stride - 2×2)

Convolutional Layer 2(Activation - Rectified Linear Unit; Filter - 12×12×60×30; Bias - 30; Stride - 1×1)

Max-pool Layer 2(Filter - 2×2; Stride - 2×2)

Convolutional Layer 3(Activation - Rectified Linear Unit; Filter - 6×6×30×15; Bias - 15; Stride - 1×1)

Max-pool Layer 3(Filter - 2×2; Stride - 2×2)

Fully Connected Layer 1(Activation - Rectified Linear Unit; Weights - 3840×4096; Bias - 4096)

Fully Connected Layer 2(Activation - Rectified Linear Unit; Weights - 4096×256; Bias 256)

Fully Connected Layer 3(Activation - Sigmoid; Weights - 256×1; Bias 4096)

2. *Recognition of partially detected probe*

Although the existing system has the advantage of cost-effectiveness, it has severe disadvantages like slow image processing, requiring high memory in performance.

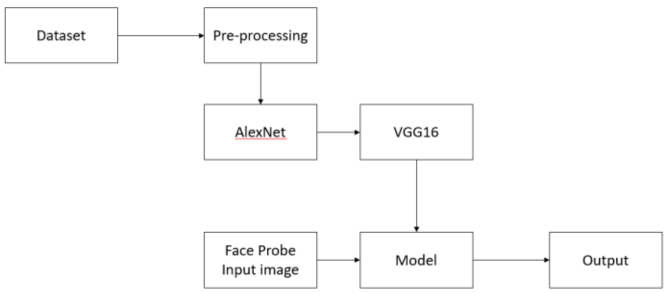

Figure 6. The Architecture of face recognition

Standard recognition networks demand fixed-sized input data and create non-spatial outputs. The fully-connected layers are used for classification and retrieval tasks compute feature representation of fixed size, ignoring spatial coordinates. CNN, with fully connected layers, cannot understand feature representation from inputs of arbitrary-sized with the length of the resulting feature vector is required to be a pre-defined fixed size.

The probe is obtained from partial face detection algorithm and passed to the face recognition model which includes VGG16 face recognition model. The layers are fine tuned as the figure 7.

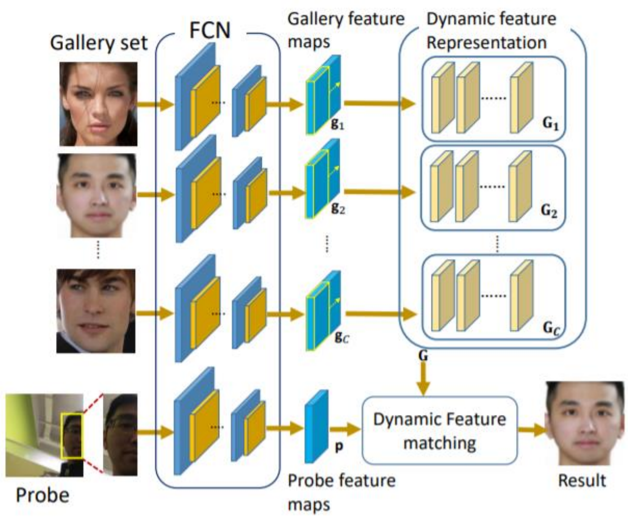

Figure 7. The Flowchart of partial face recognition

type	input	conv	conv	mpool	conv
name	-	conv1-1	conv1-2	pool1	conv2-1
support	-	3	3	2	3
filt dim	-	3	64	-	64
num filts	-	64	64	-	128
stride	-	1	1	1	1
pad	-	1	1	0	1
type	conv	mpool	conv	conv	conv
name	conv2-2	pool2	conv3-1	conv3-2	conv3-3
support	3	2	3	3	3
filt dim	128	-	128	256	256
num filts	128	-	256	256	256
stride	1	2	1	1	1
pad	1	0	1	1	1
type	mpool	conv	conv	conv	mpool
name	pool3	conv4-1	conv4-2	conv4-3	pool4
support	2	3	3	3	2
filt dim	-	256	512	512	-
num filts	-	512	512	512	-
stride	2	1	1	1	2
pad	0	1	1	1	0
type	conv	conv	conv	mpool	
name	conv5-1	conv5-2	conv5-3	pool5	
support	3	3	3	2	
filt dim	512	512	512	-	
num filts	512	512	512	-	
stride	1	1	1	2	
pad	1	1	1	0	

Figure 8. Detailed structure of our model layers for face recognition

One of the significant challenges which we encountered while training the network was the problem of minimal weight values. For training, we implemented Adam Optimizer of TensorFlow. We had practiced the cross-entropy loss function simultaneously with a regularization parameter of 0.01. We remarked that although the cost is decreasing, the training efficiency is made at 80%.

The loss function practiced was not efficient at all, as predicting all negatives in every batch will lead to this accuracy. As assumed, this gave a comparable efficiency of 80% in our test-set also. By Multiplying the positive error value with a constant > 3.5, we modified the cost function slightly. Such a modification increases the training efficiency of our method, and we obtain accuracies of >= 85%. We saved each of the models and ran it on our test samples.

IV. Conclusion

We have proposed an approach for partial face detection and recognition. Our approach shows that already available partial face recognition methods like I2C, RPSM, MKDSRC-GTP, SFM, MR-CNN, and act lower to the proposed method DFM. This is as expected because these techniques accomplish partial face detection and recognition using limited features, which easily result in mismatching and not robust. Few deep learning algorithms consider that the person's full-face image is available for matching. If this consideration is invalid, there is no algorithm in deep learning models that can work effectively with the difficulties of matching a partial face. The proposed method has exhibited promising results on the simulated partial face database (Partial-LFW). This approach would help to detect and to recognize partial face detection with the low computational cost.

References

[1] M. Yang and L. Zhang. Gabor feature based sparse representation for face recognition with gabor occlusion dictionary. European Conference on Computer Vision (ECCV), pages 448–461, 2010.

[2] Y. Sun, X. Wang, and X. Tang. Deep learning face representation from predicting 10,000 classes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1891–1898, 2014.

[3] L. He, H. Li, Q. Zhang, Z. Sun, and Z. He. Multiscale representation for partial face recognition under near infrared illumination. In IEEE International Conference on Biometrics Theory, Applications and Systems (BTAS), 2016.

[4] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1701–1708, 2014.

[5] O. M. Parkhi, A. Vedaldi, A. Zisserman, et al. Deep face recognition. In BMVC, pages 1–12, 2015

[6] W.-S. Zheng, X. Li, T. Xiang, S. Liao, J. Lai, and S. Gong. Partial person re-identification. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 4678–4686, 2015.

[7] R. Weng, J. Lu, and Y.-P. Tan. Robust point set matching for partial face recognition. IEEE Transactions on Image Processing (TIP), 25(3):1163– 1176, 2016.

[8] Y. Xu, D. Zhang, J. Yang, and J.-Y. Yang. A two-phase test sample sparse representation method for use with face recognition. IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), 21(9):1255– 1262, 2011.

[9] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 31(2):210–227, 2009.

[10] D. Yi, Z. Lei, S. Liao, and S. Z. Li. Learning face representation from scratch. arXiv preprint arXiv:1411.7923, 2014.

[11] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. In European conference on computer vision (ECCV), pages 818–833. Springer, 2014.

[12] Lenc, L.; Král, P. Automatic face recognition system based on the SIFT features. Comput. Electr. Eng. 2015.

[13] Alahi, A.; Ortiz, R.; Vandergheynst, P. Freak: Fast retina keypoint. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012.

[14] Ghorbel, A.; Tajouri, I.; Aydi, W.; Masmoudi, N. A comparative study of GOM, uLBP, VLC and fractional Eigenfaces for face recognition. In Proceedings of the 2016 International Image Processing, Applications and Systems (IPAS), Hammamet, Tunisia, 5–7 November 2016; IEEE: Piscataway, NJ, USA, 2016.

[15] Agarwal, R.; Jain, R.; Regunathan, R.; Kumar, C.P. Automatic Attendance System Using Face Recognition Technique. In Proceedings of the 2nd International Conference on Data Engineering and Communication Technology; Springer: Singapore, 2019.

[16] Lingxiao He, Haiqing Li, Qi Zhang, Zhenan Sun Member. Dynamic Feature Matching for Partial Face Recognition. In Proceedings of the IEEE Conference on Computer Vision (ISSV).

[17] Khan, S.A.; Ishtiaq, M.; Nazir, M.; Shaheen, M. Face recognition under varying expressions and illumination using particle swarm optimization. J. Comput. Sci. 2018.

[18] Vinay, A.; Cholin, A.S.; Bhat, A.D.; Murthy, K.B.; Natarajan, S. An Efficient ORB based Face Recognition framework for Human-Robot Interaction. Procedia Comput. Sci. 2018.

[19] Yang, W.J.; Chen, Y.C.; Chung, P.C.; Yang, J.F. Multi-feature shape regression for face alignment. EURASIP J. Adv. Signal Process. 2018, 2018, 51.

[20] Ouanan, H.; Ouanan, M.; Aksasse, B. Non-linear dictionary representation of deep features for face recognition from a single sample per person. Procedia Comput. Sci. 2018.

[21] Fathima, A.A.; Ajitha, S.; Vaidehi, V.; Hemalatha, M.; Karthigaiveni, R.; Kumar, R. Hybrid approach for face recognition combining Gabor Wavelet and Linear Discriminant Analysis. In Proceedings of the 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), Bhubaneswar, India, 2–3 November 2015; IEEE: Piscataway, NJ, USA, 2015.