https://doi.org/10.2352/ISSN.2470-1173.2020.6.IRIACV-048

A Review and Quantitative Evaluation of Small Face Detectors in Deep Learning

Hua Wu, Shuang Yang, Weihua Xiong, Shanhu Yu,Xinsheng Sun, and Tongqi Wei; Eaglesens Inc, Cupertino, California, USA

Abstract
Face detection is crucial to computer vision and many similar applications. Past decades have witnessed great progress in solving this problem. Contrary to traditional methods, recently many researchers have proposed a variety of CNN(Convolutional Neural Network) methods and have given out impressive results in diverse ways. Although many comprehensive evaluations or reviews about face detection are available, very few focuses on small face detection strategies. In this paper, we systematically survey some of the prevailing methods; divide them into two categories and compare them qualitatively on three real-world image data sets in terms of mAP. The experimental results show that feature pyramid with multiple predictors can produce better performance, which is helpful in future direction of research work.

Introduction
Face detection is an essential and important problem in computer vision, since it is a cornerstone of many successful applications including, but not limited to, face verification and recognition, people counting, facial shape reconstruction, face tagging and retrieval, as well as organization and presentation of digital photo-albums. The objective of face detection is to determine if there are any faces in the image and, if present, return the position of each face.

Automatic face feature and face detection have a long history that dates back to works 40-50 years ago [1-10]. With the great achievement of deep Convolutional Neural Network (CNN) on image classification in 2012, face detection has entered a new era[11-13]. Different from those traditional methods using hand-crafted features[4-10], CNN learned discriminative features through a hierarchy of non-linear information or concepts and integrate these features into a classifier. CNN has achieved impressive results on the large and medium faces in the last couple of years; however, it must face another big challenge on those small faces: normally, face size is smaller than 400 pixels (e.g. 20×20). There are three major difficulties in detecting small faces [14,15,35]: (1)These faces lack sufficient information that distinguish it from the background; (2) A several convolution and pooling layers in the CNN, the feature map will include more unnecessary background information into small faces, making it harder to detect them accurately; (3) The existing anchor strategies easily mismatches small faces' receptive fields.

Traditionally, there are potentially two ways to improve the accuracy of small faces detection. One way is to build image pyramids and try to enhance small faces' details [14,16]; the other is to build feature pyramids and consider contexts by integrating different feature levels[15]. Recently, more researchers have focused on improving the accuracy of small faces detection that claim to be superior to existing individual ones [14-37]. Therefore, the concern about the performance comparison between the competing methods makes a comparative evaluation a necessity.

Related Work
Although there are many publications on comprehensive reviews of face detection and face recognition [38-44], they only work on medium- and large-face. Cha Zhang surveys several boosting-based face detection schemes and pays more attention to the various techniques used in feature extraction and optimization strategies[39]. Chauhan[41] reviews some methods and evaluate them in term of key evaluation parameters in the authentication and identification applications, but they did not give out any comparative analysis. Another survey paper presented by Stefanos Zafeiriou [1] categorizes the up-to-date face detectors into two major schemes: rigid template and learning with boosting or deep neural network and describes several representative methods in detail. The most recent face detection survey describes 8 different state-of-the-art CNN architectures that are suitable for face detection and examine their applications to video tracking on several challenging database in terms of accuracy and speed[40].

To our best knowledge, no evaluation work has been concentrated on small face detection. In this paper, we aim to fill in this gap and present a brief survey on the latest development in small face detector and present a quantitative comparison among some prevailing methods using real image data sets.

The contributions of this paper can be summarized as follows:

- It reviews and classifies several prevailing small face detection methods into two major categories based on their underlying strategies.
- It provides a comprehensive comparison of 6 representative methods on 3 real-world image sets. The proposed comparison is valuable in analyzing the current trends in small face detection research.

Review of Several Prevailing Small Face Detectors
The existing small face detectors[14-37] vary significantly in underlying strategy. In this paper, we roughly divide them into two major categories: Image Pyramid(IP)[14,16,22,23] aims to enhance small faces' details by increasing the resolution of the whole image or potential faces; Feature Pyramid(FP)[15] tries to work on different layers of feature map and integrate them to include more contexture information into detection. FP can be further divided into 2 sub-categories: A method is called Feature Pyramid Unitary Predictor (FPU) if only one prediction is made on the highest resolution level[17,18,24-27]; otherwise, it is called Feature Pyramid Multiple Predictor(FPM) if different predictions are made on several layers independently[19,20,28-31]

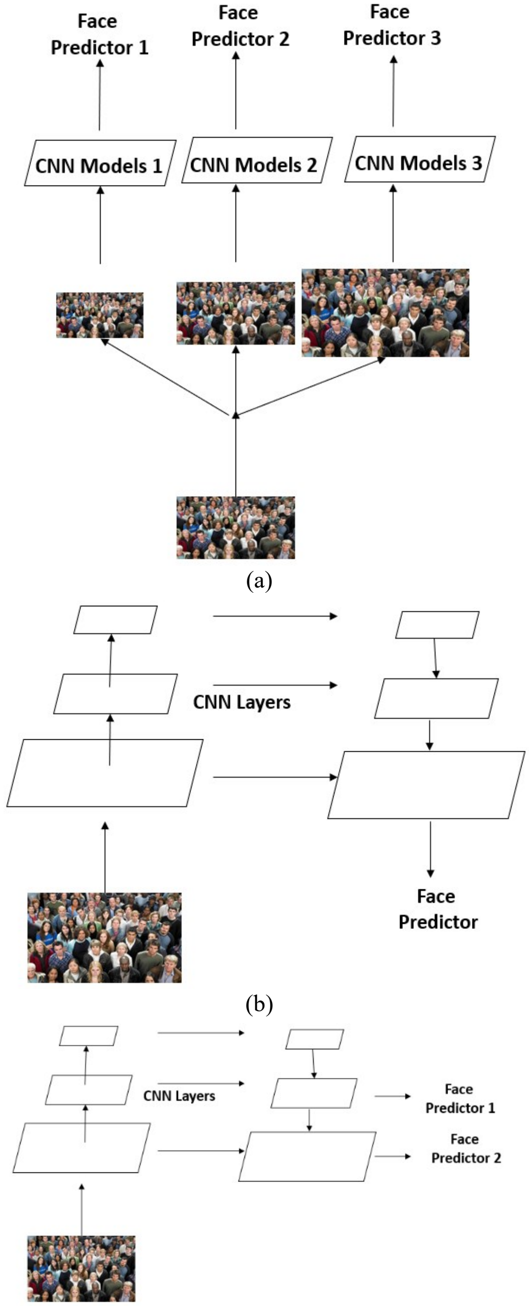


Figure 1.Three types of Small Face Detectors' Architecture, (a) Image Pyramid, (b) Pyramid Feature Unitary Predictor and (c) Pyramid Feature Multiple Predictor. (The figures refer to the relevant descriptions and legends in [15])

IP Methods

- ### Finding Tiny Faces (FTF)

A relatively earlier paper focusing on small face detection comes from [14]. Its central idea is to create image pyramids to increase the resolution of small faces, and design different detection models according on different scale images. The biggest problem with this type of method is that it is computationally complex and time consuming. Simply scaling on the full image will also add more calculations for redundant or extraneous information

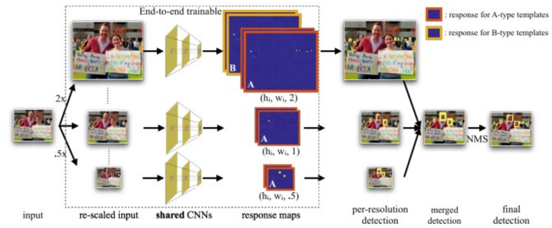


Figure 2. FTF CNN Architecture, copy from [14]

- ### Multi-task Cascaded Convolutional Networks (MTCNN)

Another type of method is to organize several sub-networks in a cascade structure. The central idea is to detect those possible areas at a relatively lower resolution, resize these potentials and determine if they are faces or not. Several representative methods are Cascade CNN[16], TG-GAN[22], SOD-MTGAN[23]. Cascade CNN[16] includes three sub-CNN Network: Proposal Network(P-Net)、Refine Network(R-Net)、Output Network(O-Net). P-Net is mainly used to generate some bounding boxes and potential faces. The inputs of R-Net are from P-Net in which a large number of non-face frames will be removed. All candidate faces, including small ones will be resized into 24*24. The final stage, O-Net, has similar function as that of R-Net. But this step will add the return of the landmark position. The input size is adjusted to be 48*48, and the output contains each face's location and confident score. The disadvantage of Cascade CNN is that it uses simple bilinear or bicubic interpolation to resize smaller faces that will lose their details.

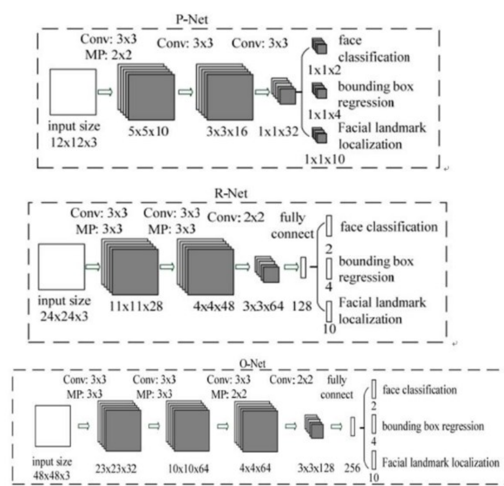


Figure 3. MTCNN Architecture, copy from [16]

PFU Method

- ### FaceMagnet(FaceMag)

FaceMagenet follows two-stage object detection's framework[17], but it defines two individual set of deconvolution layers, one in Region Proposal Network(RPN) and one before ROI Network, to improve small faces' performance[17]. It considers context information around small face by merging the features from those pre-defined original anchors and those from enlarged windows. It also adds a special branch for small face detector, since the feature maps from lower layers are more suitable for obtaining small faces. Its performance is heavily dependent on the anchor boxes definition and tilting.

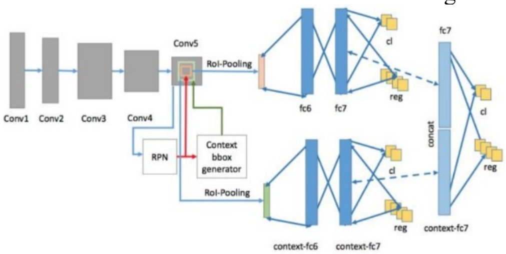


Figure 4. FaceMagnet CNN Architecture, copy from [17]

- ***Learning Small Faces on Hard Images(LFH)***

LFH builds bottom-top pathway by bicubically interpolating feature maps from higher layers and concatenates the upsampled feature maps from different layers[18]. The concatenation result is then applied by a 3×3 convolution to create final detection feature map, which will be fed into the detection head for classification and bounding-box regression. However, its performance is limited with only one skip-connection between one upsampled layer and corresponding higher layer.

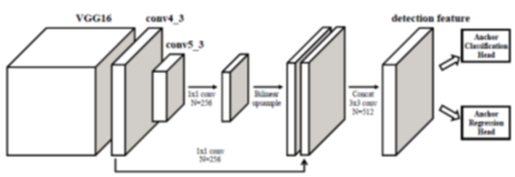


Figure 5. LFH CNN Architecture, copy from [18]

PFM Method

- ***Dual Shot Face Detector (DSFD)***

Dual Shot Face Detector not only aggregates high-level and low-level output feature maps, but also considers the information of the current layer[19]. In DSFD, there are two hierarchical feature extraction modules, and each of the two modules has a corresponding loss function and face detector. The big shortcoming is its parameter size, doubling that from normal single shot face detector.

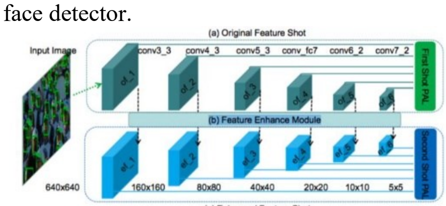


Figure 6. DSFD CNN Architecture, copy from [19]

- ***Extreme Tiny Face Detector via Iterative Filter Reuse(EXTD)***

EXTD combines the advantages of SSD[21] and FPN[15] that enable low-level feature maps to capture sufficient target semantic information[20]. The basic approach is to iteratively use a shared lightweight shallow backbone network to generate each feature map. The feature implies that the its inference time is still longer, although its parameter size is relatively smaller.

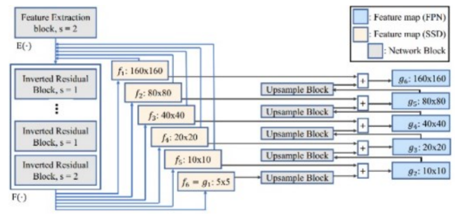


Figure 7. EXTD CNN Architecture, copy from [15]

Table1 gives out a summary of pros and cons of these six methods used in the evaluation in the following experiment.

Table 1 Pros and Cons of Six Small Face Detectors

Cate-gory	Method Name	Pro	Con
IP	FTF	Different CNN on different scale of image	Higher computational complexity
	MTCNN	Inference is faster	Easily lose details
FPU	FaceMag	Special branch for small face	Depends on anchor boxes heavily
	LFH	Structure is simpler	Network is not deep enough
FPM	DSFD	Enhance every feature map	Too many parameters
	EXTD	Less parameter size	Longer Inference Time

Experiment Setup

These methods are tested on 3 different real-world image database and the overall performance of face detection, especially small face detection, are compared. We downloaded the model files and parameters of FTF, MTCNN,LFH, DSFD,EXTD from Websites[45], and trained FaceMag's model using WiderFace Training Data Set[46]. Their parameter sizes are summarized in the following table.

Table 2 Six Small Face Detectors

Category	Method Name	Parameter Size(MB)
IP	FTF	98.93
	MTCNN	2.27
FPU	FaceMag	540.46
	LFH	71.19
FPM	DSFD	469.73
	EXTD	0.681

WiderFace Database

Among many databases used for face detection evaluation[47], WiderFace has received the most attention in recent years[46]. The database has a total of 32,203 images, including 393,703 faces, and there are great changes in the size, posture, occlusion, expression, makeup, and illumination of the face.

3226 images from this database are defined as evaluation data set. There are 39,708 faces in the set, but only 31,958 valid faces are often used in performance evaluation, the remaining

8000 faces are excluded and generally not considered due to lower resolution or image quality, we called them 'Invalid Faces'. According to the detection rate of EdgeBox[48], all valid faces are divided into three difficulty subsets: Easy, Medium, Hard.

However, we note that the evaluation and comparison of the small face detection method should ultimately be based on the face size rather than difficulty levels, since each difficulty subset contains faces of different sizes. Therefore, we reclassified these valid face images into 3 subsets based on their areas. Those faces with an area greater than $1600(40 \times 40)$ are categorized into 'Large' subset, the faces with an area between $400(20 \times 20)$ and $1600(40 \times 40)$ belongs to 'Normal' subset, the face with an area smaller than 400 (20×20) is categorized into 'Small' subset. As expected, it is quite challenging to achieve good detection performance on the Small subset.

Moreover, we notice that, although most of those excluded invalid face images are small ones with height or width smaller than 10 pixels, exact detection of them is also very important in some applications (e.g. people counting, people density analysis). Figure 8 gives out an example in which the faces marked with blue rectangle are valid ones with number of 17 and faces marked with green are invalid ones with number of 566. Therefore, we will also include these invalid ones and evaluate the detectors on all face images in terms of area size.

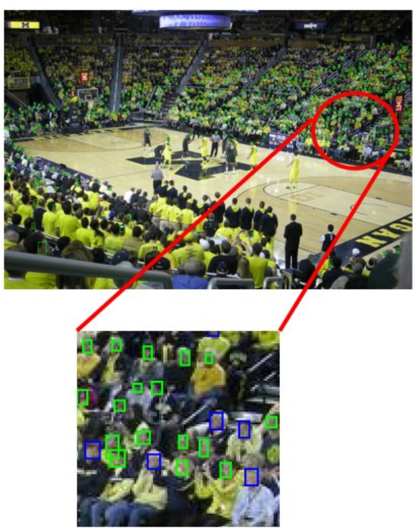


Figure 8. An example from WiderFace Evaluation Dataset has 583 labelled faces, only 17 of them are Valid Ones (Shown in Blue), and the other 566 faces (shown in Green) are excluded in regular face detector evaluations. However, accurate detection of these excluded faces is also very important in some real applications.

The following table gives out the face number from different subset. The 3 columns in 'Valid Dataset' field shows face number from two different perspectives: difficulty level and face area. The 5th column is the invalid face number in terms of face area. The last column shows the total face number having different area, each total number is the sum of Valid Hard face number and that from invalid dataset, since Easy/Medium are subsets of Hard.

Table 3. The face number statistics of WiderFace Database

	Valid Dataset			Invalid Dataset	Total
	Easy	Medium	Hard		
Small	303	1146	14721	7515	22136
Normal	837	4704	7469	261	9339
Large	6071	9078	8159	77	8233

FDDB Dataset

Face Detection Data Set and Benchmark (FDDB) [49] is another popular face detection evaluation platform, with 2,845 images and a total of 5,171 faces. It has gray and color images with different poses, different resolutions, rotation and occlusion. The standard face labeling area is elliptical, thus we relabeled their locations via using a rectangle to bound the ellipse.

UFDD Dataset

The group of researchers headed by Hajime Nada from Fujitsu published an Un-constrained Face Detection Dataset(UFDD) in 2018 [50]. Images were collected from different sources on the web, such as Google, Bing, Yahoo, Creative commons search, Pixabay, Pixels, Wikimedia commons, Flickr, Unsplash, Vimeo, and Baidu. It includes 6,424 images with 10,895 annotations and captures variations in weather conditions (rain, snow, haze), motion and focus blur, illumination variations, lens impediments.

Of note is that the UFDD dataset also includes a large set of distractor images that is usually ignored by the existing datasets. Distractors either contain non-human faces such as animal faces or no faces at all. The presence of such images is especially important to measure the performance of a face detector in rejecting non-face images and to study the false positive rate of the algorithms.

Table 4. The face number statistics of FDDB and UFDD Database

Database	Small	Normal	Large
FDDB	56	306	4809
UFDD	2574	4108	4213

Experiment and Analysis

In this section, we will test and analyze 6 selected small face detectors on 3 face databases. Performance is evaluated in terms of mAP. Since we pay more attention to the performance of small face detectors, we highlight the results from 'Hard' subsets of WiderFace and those from 'Small' subset.

Results on the WiderFace Database

The first experiment is carried on WiderFace database. Table 5-7 show the mAP results on it. It is not difficult to find that if we compare small face detection methods on the Hard data set in terms of difficulty level shown in Table 5, the best performance is from two methods belonging to FPM, which can reach 0.820 and 0.844 respectively. In particularly, although EXTD is not the best one, its performance is still impressive since its parameter size is much smaller than others. The FTF from IP also shows competitive performance with mAP of 0.802, but the image pyramid will significantly increase the computational cost and have lower inference speed. The performance from two methods of FPU is not bad，which can have 0.757 or 0.771 individuals. The performance of MTCNN is only 0.607， which seems to indicate that only increasing face resolution is not suitable for small face detection.

Table 5. Comparison of mAP Results of 6 small face detectors on WiderFace Valid Faces in terms of difficulty levels

Category	Method	Easy	Medium	Hard
IP	FTF	0.869	0.809	***0.802***
	MTCNN	0.851	0.820	***0.607***
FPU	FaceMag	0.892	0.872	***0.757***
	LFH	0.866	0.786	***0.771***
FPM	DSFD	0.945	0.931	***0.844***
	EXTD	0.912	0.899	***0.820***

Table 6. Comparison of mAP Results of 6 small face detectors on WiderFace Valid Faces in terms of face area

Category	Method	Large	Normal	Small
IP	FTF	0.874	0.829	***0.695***
	MTCNN	0.889	0.80	***0.665***
FPU	FaceMag	0.918	0.879	***0.519***
	LFH	0.825	0.806	***0.566***
FPM	DSFD	0.931	0.839	***0.660***
	EXTD	0.919	0.823	***0.628***

Table 7. Comparison of mAP Results of 6 small face detectors on WiderFace All Faces in terms of face area

Category	Method	Large	Normal	Small
IP	FTF	0.899	0.865	***0.718***
	MTCNN	0.898	0.766	***0.667***
FPU	FaceMag	0.935	0.825	***0.605***
	LFH	0.847	0.803	***0.604***
FPM	DSFD	0.945	0.884	***0.747***
	EXTD	0.945	0.870	***0.669***

However, if we review the mAP results from the perspective of face area, the actual situation may vary. We can notice that whether we look at the outputs from valid faces shown in Table 6 or from all the faces shown in Table 6, the performance of most methods has inevitably degraded to some degree. For example, the performance of DSFD and EXTD on valid faces have decreased by 21.8% and 23.4 %, reaching to 0.660 and 0.628; the FTF method came to 0.695; the mAP of two FPU methods dropped below 0.6, which is the worst among all comparison methods. Only the performance of MTCNN increases, and its comprehensive mAP value is comparable to those of FPM/FTF methods. This is likely due to MTCNN performing the appropriate upsampling operation on the small face.

Therefore, we can suggest that, in order to ensure the performance of Normal/Large face detection as well as improve the performance of small faces, a good potential way is to integrate face resolution enhancement and feature pyramid with multiple predictors. And it is not recommended to adopt feature pyramid strategy with unitary prediction.

Results on the FDDB Database

Next, we consider performance comparison on FDDB database. Table 8 give out the experimental results. The overall performance on Normal and Small dataset is worse than that from WiderFace Database. This is reasonable since all these methods' models are trained on the WiderFace Training Dataset. For the comparison and analysis results among different methods on this database, similar conclusions as those on WiderFace can be gained.

Table 8. Comparison of mAP Results of 6 small face detectors on FDDB All Faces in terms of face area

Category	Method	Large	Normal	Small
IP	FTF	0.892	0.735	***0.633***
	MTCNN	0.881	0.745	***0.604***
FPU	FaceMag	0.916	0.728	***0.548***
	LFH	0.802	0.787	***0.540***
FPM	DSFD	0.914	0.701	***0.635***
	EXTD	0.935	0.780	***0.614***

Results on the UFDD Database

Our final test with real-world face images is based on UFDD Database. Table 9 give out the experimental results. These results tell us almost the same story as that from experiment on WiderFace and FDDB.

Table 9. Comparison of mAP Results of 6 small face detectors on UFDD All Faces in terms of face area

Category	Method	Large	Normal	Small
IP	FTF	0.880	0.791	***0.623***
	MTCNN	0.868	0.781	***0.641***
FPU	FaceMag	0.893	0.764	***0.489***
	LFH	0.894	0.819	***0.437***
FPM	DSFD	0.899	0.773	***0.619***
	EXTD	0.904	0.772	***0.610***

Conclusion

Reviewing the comparison among all methods in each category, we can find that find the following points：(1)All methods have similar performances on large and normal face detections; (2) The overall performances of DSFD and EXTD in FPM category and those of FTF and MFCNN in IP category are better than those in FPU on small face detectors, showing that multiple predictors are necessary to improve small face detection. Understanding such factors will help potential directions for future research.

Acknowledge

We thank Jeffrey Xiong for proofreading the manuscript.

Reference

[1] S. Zafeiriou, C. Zhang, and Z. Zhang, "A survey on face detection in the wild: past, present and future", Computer Vision and Image Understanding, Vol 138, pp1–24, 2015.

[2] M.-H. Yang, D. J. Kriegman, N. Ahuja, "Detecting faces in images: A survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 24, no 1, pp 34–58, 2002

[3] E. Hjelmas, B. K. Low, “Face detection: A survey”, Computer Vision and Image Understanding, Vol 83, pp 236–274,2001

[4] P. Viola and M. J. Jones, “Robust real-time face detection”, International Journal of Computer Vision, Vol 57, no 2, pp137–154, 2004

[5] P Dollar, R. Appel, S. Belongie, and P. Perona. “Fast Feature Pyramids for Object Detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(8): 1532–1545, 2014

[6] E. Ohn-Bar, and M. Trivedi, “Fast and Robust Object Detection Using Visual Subcategories”, In the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2014.

[7] T. Ahonen, A. Hadid, and M. Pietikainen, “Face Description with Local Binary Patterns: Application to Face Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(12): 2037–2041, 2006

[8] S.G. Tong, Y.Y. Huang, Z.M. Tong, “A Robust Face Recognition Method Combining LBP with Multi-mirror Symmetry for Images with Various Face Interferences”, International Journal of Automation and Computing, Vol. 16, Issue 5, pp671-682, 2019

[9] J.Yan, X.Zhang, Z.Lei,,and S.Z.Li, “Real-time high performance deformable model for face detection in the wild”, International Conference on Biometrics,2013

[10] J. Yan, Z. Lei, L. Wen, and S. Z. Li, “The fastest deformable part model for object detection”, IEEE Conference on Computer Vision and Pattern Recognition, 2014

[11] Z.Q. Zhao, P. Zheng, S.t. Xu, X. Wu, ”Object Detection with Deep Learning: A Review”, IEEE Conference on Computer Vision and Pattern Recognition, 2019

[12] C. Zhang, X. Xu, and D. Tu, “Face Detection using Improved Faster Rcnn”, IEEE International Conference on Automatic Face & Gesture Recognition (FG),2015

[13].A. Kalinovskiy, V.G. Spitsyn, “Compact Convolutional Neural Network Cascade for Face Detection”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2015

[14] P. Hu and D. Ramanan, “Finding tiny faces”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017

[15] T-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature pyramid networks for object detection”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017,

[16] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao. “Joint face detection and alignment using multitask cascaded convolutional networks”, Signal Processing Letter, Vol 23, no 10, pp1499–1503, 2016.

[17] P. Samangouei, M. Najibi, L. Davis, R. Chellappa, “Face-MagNet: Magnifying Feature Maps to Detect Small Faces”, IEEE Winter Conference on Applications of Computer Vision,2018

[18] Z Zhang, W Shen, S Qiao, Y Wang, B Wang, A Yuille, “Robust Face Detection via Learning Small Faces on Hard Images”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018

[19] J. Li, Y. Wang, C. Wang, Y. Tai, J. Qian, J. Yang, C. Wang, J. Li, and F. Huang. “Dsfd: Dual shot face detector”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2019.

[20] Y.j. Yoo, D. Han, S. Yun, “EXTD: Extremely Tiny Face Detector via Iterative Filter Reuse”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2019.

[21] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, “SSD: Single shot multibox detector“. IEEE European Conference on Computer Vision(ECCV), 2016

[22] Y. Bai, Y. Zhang, M. Ding, and B. Ghanem, “Finding Tiny Faces in the Wild with Generative Adversarial Network”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018

[23] Y. Bai, Y. Zhang, M. Ding, and B. Ghanem, “SOD-MTGAN: Small Object Detection via Multi-Task Generative Adversarial Network”, IEEE Conference on European Conference on Computer Vision (ECCV), 2018

[24] S. Bell, C. Zitnick,K. Bala, and R. Girshick, “Inside-Outside Net: Detecting Objects in Context with Skip Pooling and Recurrent Neural Networks”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018

[25] W. Liu, S. Liao, W. Ren, W. Hu, and Y. Yu, “CSP: Center and Scale Prediction”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019

[26] H.Zhao, J. Shi, Z. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017

[27] C. Zhu, Y. Zheng,K. Luu, and M. Savvides, “CMS-RCNN: Contextual Multi-Scale Region-based CNN for Unconstrained Face Detection”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

[28] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, “S3FD: Single Shot Scale-invariant Face Detector”, IEEE Conference on Computer Vision and Pattern Recognition,(CVPR) 2017

[29] M. Najibi, P. Samangouei, R. Chellappa, and L. Davis, “SSH: Single Stage Headless Face Detector”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017

[30] G. Cao, X. Xie, W. Yang, Q. Liao, G. Shi, and J. Wu, “Feature-Fused SSD: Fast Detection for Small Objects”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018

[31] X. Tang, D. K. Du, Z. He, and J. Liu, “PyramidBox: A Context-assisted Single Shot Face Detector”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018

[32] W.Xiang, D.Q. Zhang, H Yu, V. Athitsos, “Context-Aware Single-Shot Detector”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018

[33] S. Zagoruyko, A. Lerer, T.y. Lin, P.O. Pinheiro, S. Gross, S. Chintala, P. Dollar, “A MultiPath Network for Object Detection”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016

[34] Y. Xi, J. Zheng, X. He, W. Jia. H. Li, “Beyond Context: Exploring Semantic Similarity for Tiny Face Detection”, IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018

[35] C.Zhu, R. Tao, K. Luu, and M. Savvides, "Seeing Small Faces from Robust Anchor's Perspective", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018

[36] R. Lalonde, D. Zhang, M. Shah, "ClusterNet: Detecting Small Objects in Large Scenes by Exploiting Spatio-Temporal Information", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017

[37] C.Y. Fu, W. Liu, A. Ranga, A. Tyagi, A.C. Berg, "DSSD : Deconvolutional Single Shot Detector", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017

[38] S. Mittal, C. Shivnani, "Face Detection and Tracking: A Comparative study of two algorithms', International Journal of Computer Science & Communication, Vol 7, no 1, pp76-82, 2016

[39] C. Zhang and Z. Zhang, "A survey of recent advances in face detection", Technical Report MSR-TR-2010-66, 2010

[40] E. Granger, M. Kiran, L. Blais, "A comparison of CNN-based face and head detectors for real-time video surveillance applications". Seventh International Conference on Image Processing Theory, Tools and Applications, 2017

[41] C. Mayank and M. Sakle. "Study & Analysis of Different Face Detection Techniques", International Journal of Computer Science and Information Technologies, Vol 5, no 2 pp1615-1618, 2014

[42] L. Jiao, F. Zhang, F. Liu, S. Ynag, L. Li, Z. Feng, R. Qu, "A Survey of Deep Learning-based Object Detection", IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2019

[43] L Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, M. Pietikainen, "Deep Learning for Generic Object Detection: A Survey", IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2019

[44]M. Wang, W. Deng,"Deep Face Recognition:A Survey", IEEE Conference on Computer Vision and Pattern Recognition, 2019

[45] https://paperswithcode.com/task/face-detection/latest

[46] S. Yang, P. Luo, C.-C. Loy, and X. Tang, "Wider face: A face detection benchmark", IEEE Conference on Computer Vision and Pattern Recognition, 2016

[47] http://www.face-rec.org/databases/

[48] C. Zitnick and P. Dollar. "Edge boxes: Locating object proposals from edges", IEEE European Conference on Computer Vision, 2014

[49] http://vis-www.cs.umass.edu/fddb/index.html

[50] http://www.ufdd.info/

Author Biography

Hua Wu received his BS in artificial intelligence from Guangxi University(2012). Then he worked in Allinpay Company for two years. He studied the technology of machine learning in Wuhan University from 2014 and afforded a certificate upon completion of his PhD course from Wuhan University(2017). Then he joined in unisoc Company for one year. Since 2018, he has worked in eaglesens company focusing on small object detection with deep learning technology.

Shuang Yang received his BE in communication engineering from Jilin University(2011) and his MA in Electronic science and technology from Beijing University of Posts and Telecommunications(2017).Then he worked in Vision lab at Uni-Ubi Company. Since 2019, he has worked in eaglesens company and focused on Person Re-identification and face detection based on deep learning technology

Weihua Xiong received his master and Ph.D. from Peking University and Simon Fraser University in 1996 and 2007 respectively. After that, He has served as Staff Engineer and Senior Technical Director at OV, Qualcomm and UNISOC, and is currently VP of Engineering in Eaglesens. His research interests are image processing, AI algorithms at Edge Computing, etc.

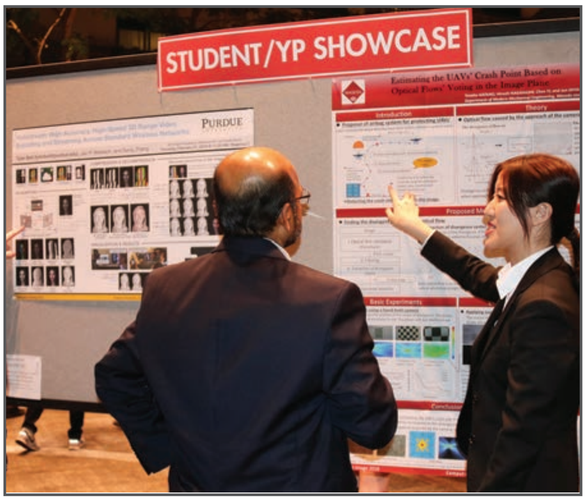
JOIN US AT THE NEXT EI!

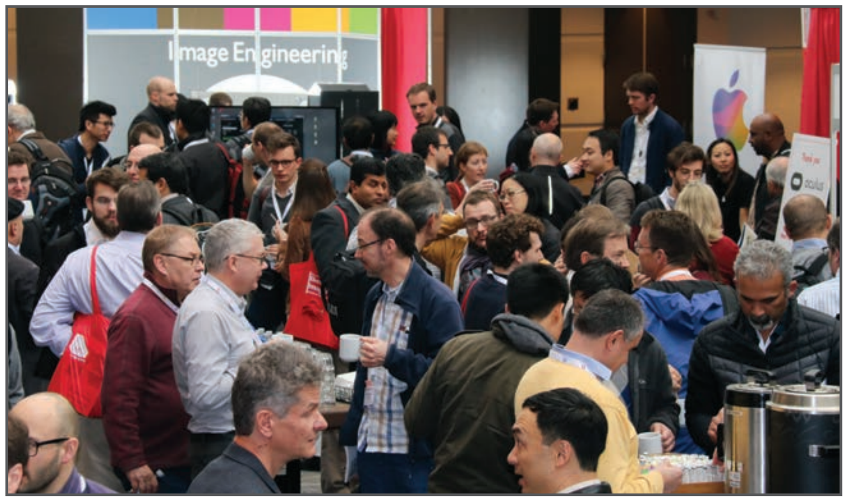
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org