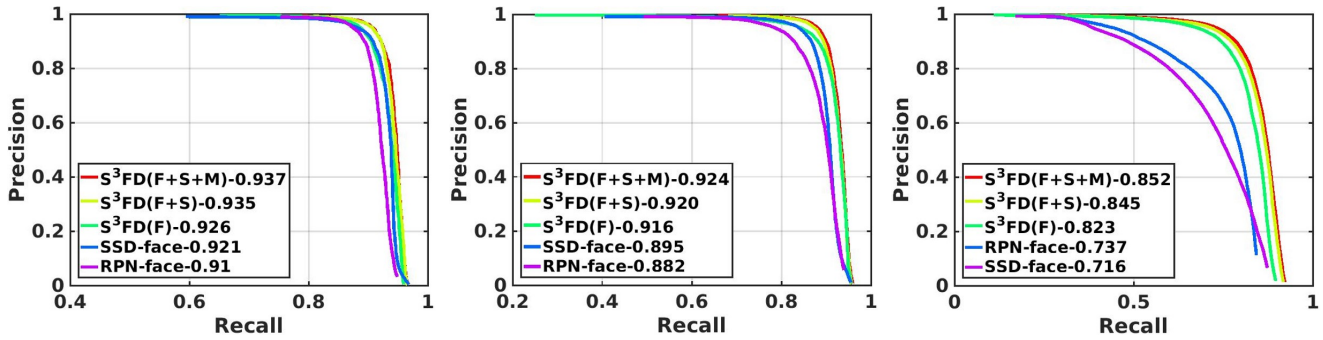
S^{3}FD: Single Shot Scale-invariant Face Detector -Supplementary Material-

Shifeng Zhang Xiangyu Zhu Zhen Lei Hailin Shi Xiaobo Wang Stan Z. Li
CBSR & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China
University of Chinese Academy of Sciences, Beijing, China
{shifeng.zhang,xiangyu.zhu,zlei,hailin.shi,xiaobo.wang,szli}@nlpr.ia.ac.cn

1. Precision-recall curves

In our submitted paper, Tab.3 in subsection 4.1 only provides the mAP of RPN-face, SSD-face, S^{3}FD(F), S^{3}FD(F+S) and S^{3}FD(F+S+M). Their precision-recall curves on the WIDER FACE validation set are shown in Fig. 1 for details.



(a) Easy (b) Medium (c) Hard

Figure 1. Precision-recall curves on WIDER FACE validation set.

2. Qualitative results

In this section, we demonstrate some qualitative results on common face detection benchmarks, including AFW (Fig. 2), PASCAL face (Fig. 3), FDDB (Fig. 4) and WIDER FACE (Fig. 5). Besides, another impressive result is shown in Fig. 6.

Figure 2. Qualitative results on AFW. The faces in these results have a high degree of variability in scale, pose and occlusion. Our S^{3}FD is able to detect these faces with a high confidence, especially for small faces. Please zoom in to see some small detections.

Figure 3. Qualitative results on PASCAL face. Most faces in these results are small faces, because the image in PASCAL face has a low resolution. Our S^{3}FD is able to handle small faces well. Please zoom in to see some small detections.

Figure 4. Qualitative results on FDDB. These results indicate that our S^{3}FD is robust to large appearance, heavy occlusion, scale variance and heavy blur. Please zoom in to see some small detections.

(a) Scale attribute. Our S^{3}FD is able to detect faces at a continuous range of scales.

(b) Our S^{3}FD is robust to pose, occlusion, expression, makeup, illumination and blur.

Figure 5. Qualitative results on WIDER FACE. We visualize some examples for each attribute. Please zoom in to see some small detections.

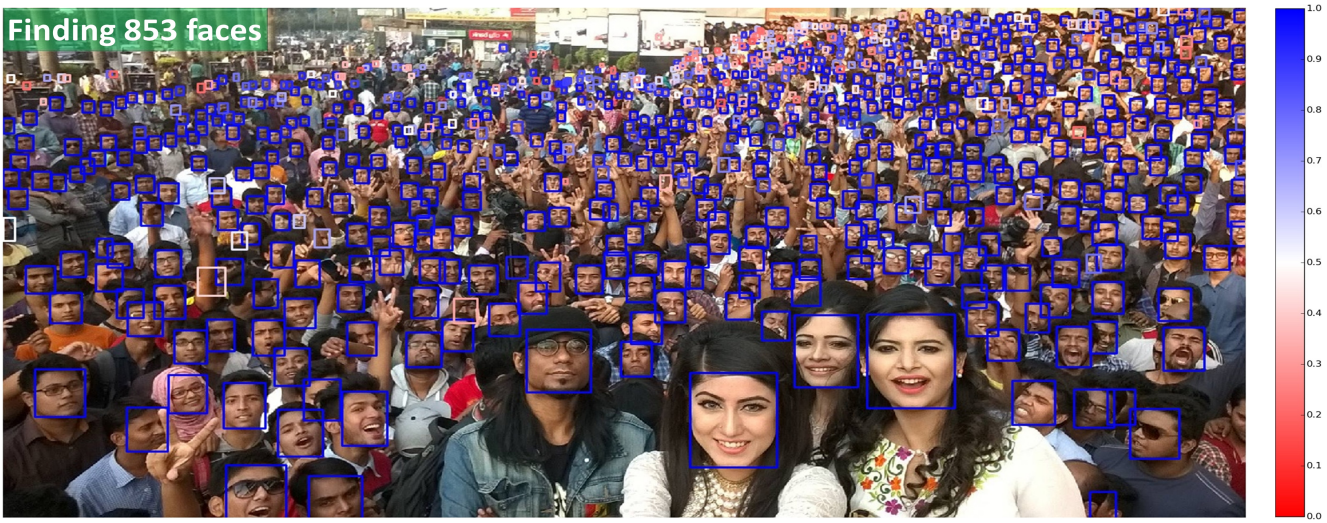


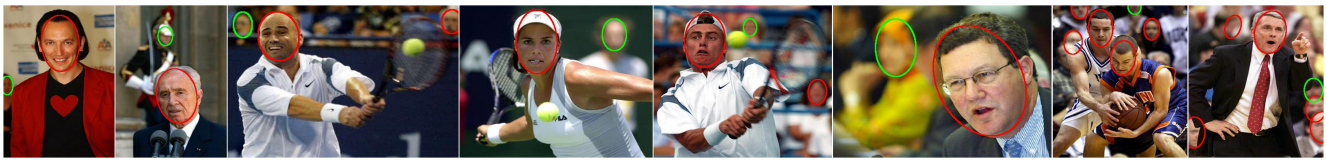
Figure 6. Another qualitative result. Our S^{3}FD can find 853 faces out of the reportedly 1000 present in the above image. Detector confidence is given by the colorbar on the right. Please zoom in to see some small detections.

3. Examples of manually labelled faces on FDDB

We add 238 unlabelled faces whose height and width are more than 20 pixels. Some examples are shown in Fig. 7.

(a) Profile faces

(b) Occluded faces



(c) Blur faces

(d) Statue faces

(e) Miscellaneous faces

Figure 7. Examples of our manually labelled faces on the FDDB dataset. Red ellipses are the faces that FDDB has already labelled, green ellipses are the newly added faces.

4. Ablative analysis of each detection layers

To examine the contribution of each detection layers on the mAP performance, we progressively remove the detection layers to test their contribution on the WIDER FACE Val set. The detailed experiment results are listed in Tab. 1. After removing Conv3_3 layer, the mAP changes are +0.3%(Easy), +0.5%(Medium) and -24.7%(Hard), showing Conv3_3 is crucial to detect small faces, but tiling plenty of smallest anchors also slightly hurts medium and large face detection performance. Besides, the most contribution of Easy and Medium subset are Conv5_3 (25.8%) and Conv4_3 (20.6%), respectively.

Detection layers	**Ablative analysis**					
Conv3_3	×					
Conv4_3		×				
Conv5_3			×			
Conv_fc7				×		
Conv6_2					×	
Conv7_2						×
mAP changes on Easy subset (%)	+0.3	-0.6	**-25.8**	-10.2	-3.2	-1.4
mAP changes on Medium subset (%)	+0.5	**-20.6**	-12.2	-5.0	-1.5	-0.7
mAP changes on Hard subset (%)	**-24.7**	-8.7	-4.1	-1.8	-0.6	-0.2

Table 1. The ablative results of each detection layers on the WIDER FACE Val set.