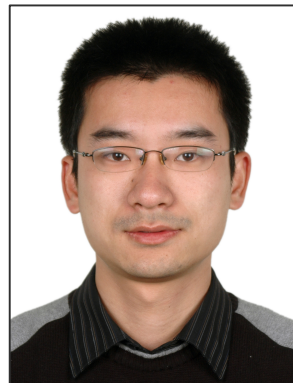# Team Member

Shifeng Zhang   Cheng Chi   Zhen Lei   Stan Z. Li

University of Chinese Academy of Sciences

Track

- **<u>Face Detection</u>**
  - Third Place Award
- Pedestrian Detection
- Person Search

Problem

- **Recall Efficiency**
  - Number of false positives needs to be reduced at the high recall rates
  - AP is very high, but precision is not high enough at high recall rates
  - Precision is only ~50% at recall rate = 90%
  - Half of detections are false positives

- **Location Accuracy**
  - Accuracy of the bounding box location needs to be Improved
  - Adopt MS COCO evaluation criterion
  - Put more emphasis on the bounding box location accuracy
  - As the IoU threshold increases, the AP drops dramatically

Solution
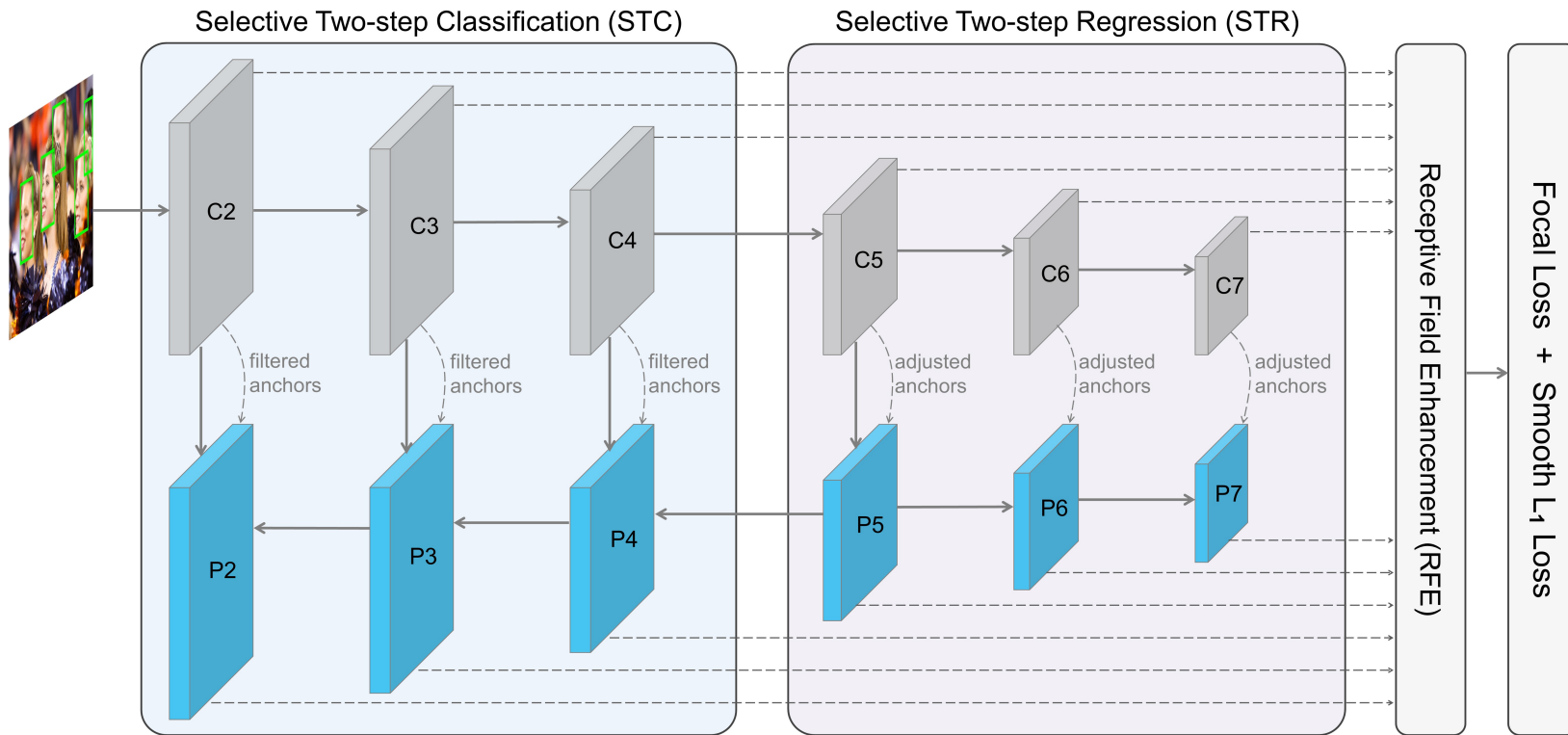
- Selective Refinement Network (SRN): STC、STR、RFE

1. Selective Two-Step Classification (STC)

- Need to tile plenty of small anchors to detect small faces

- Cause extreme class imbalance between positives and negatives

- The number of positive samples is only a few dozen or less

- Doing two-step classification is essential to reduce the false positives

- Performing two-step classification on all pyramid levels is unnecessary

STC	B	P2	P3	P4	P5	P6	P7
Easy	95.1	95.2	95.2	95.2	95.0	95.1	95.0
Medium	93.9	94.2	94.3	94.1	93.9	93.7	93.9
Hard	88.0	88.9	88.7	88.5	87.8	88.0	87.7

- Select P2, P3, and P4 to perform two-step classification

1. Selective Two-Step Classification (STC)
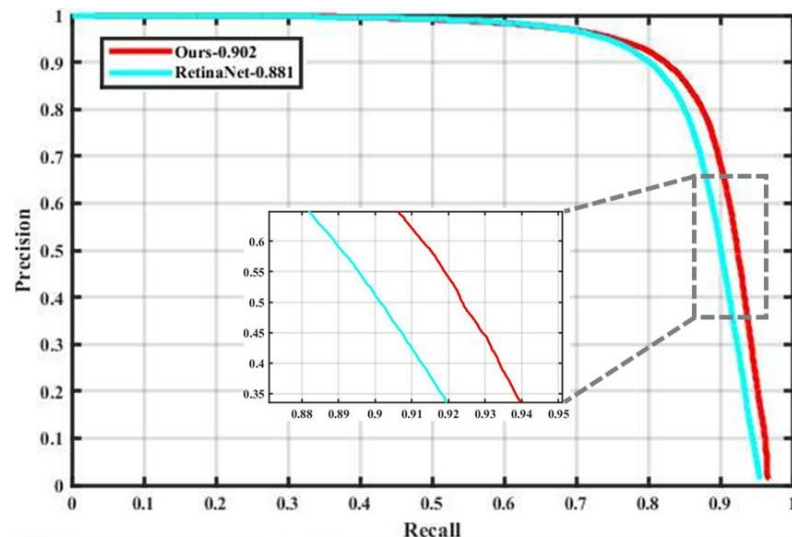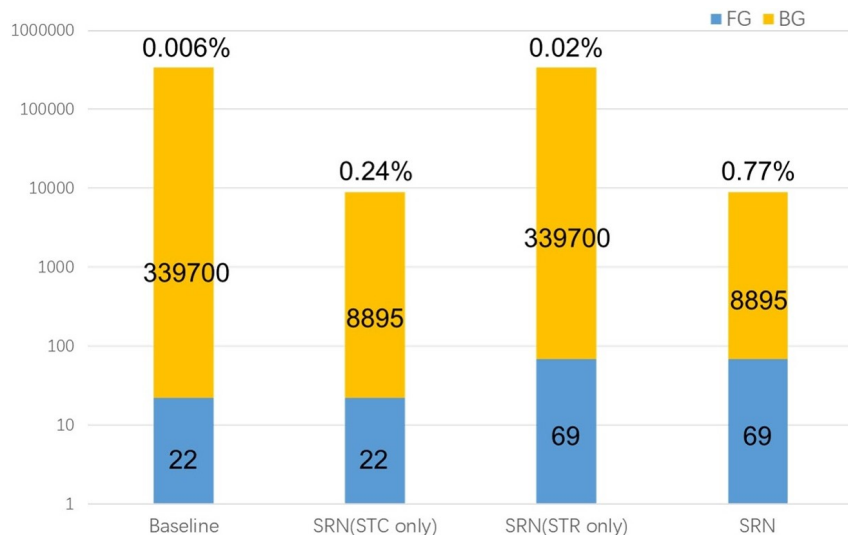
- Increase the positives/negatives ratio by about 114 times

- Improve the precision by about 20% at high recall rates

2. Selective Two-Step Regression (STR)

- Making the location of bounding box more accurate is a challenging issue

- Current one-stage methods rely on one-step regression

- It is inaccurate the in MS COCO evaluation metric

- Blindly adding multi-step regression is often counterproductive

STR	B	P2	P3	P4	P5	P6	P7
Easy	95.1	94.8	94.3	94.8	**95.4**	**95.7**	**95.6**
Medium	93.9	93.4	93.7	93.9	**94.2**	**94.4**	**94.6**
Hard	88.0	87.5	87.7	87.0	**88.2**	**88.2**	**88.4**

- Select P5, P6, and P6 to perform two-step regression

2. Selective Two-Step Regression (STR)
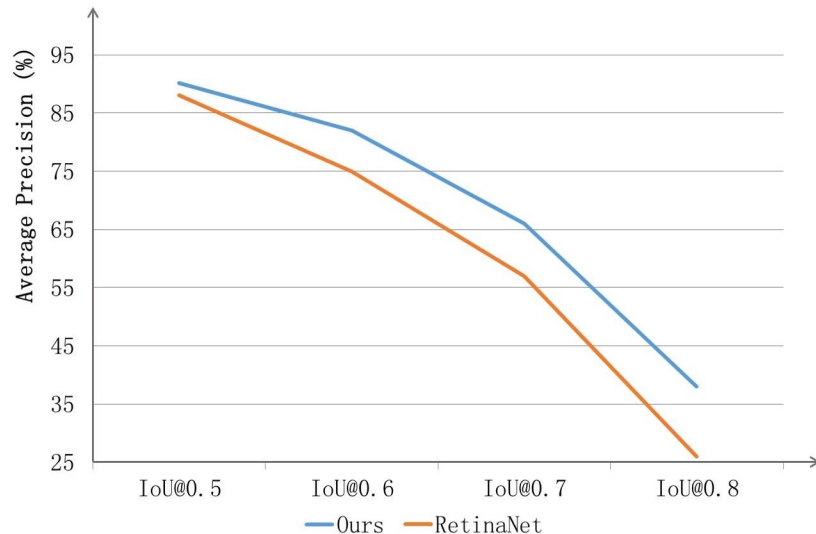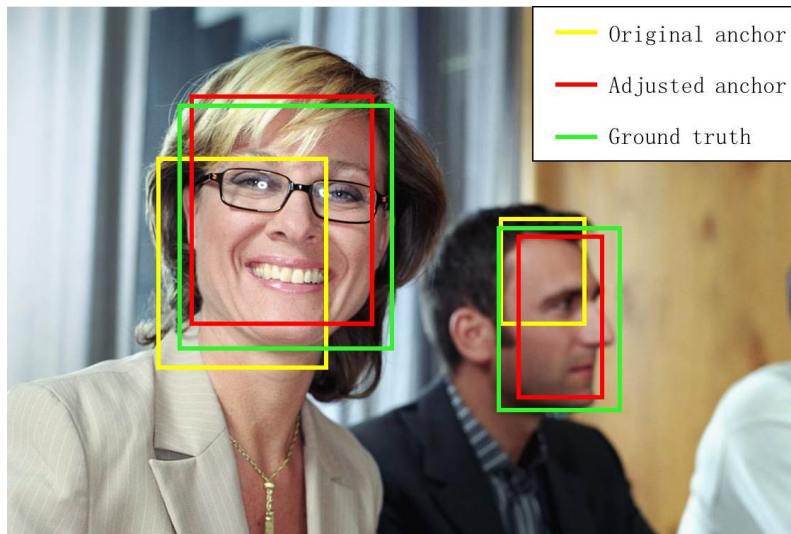
- Provide better initialization for the subsequent regressor

- Produce more accurate locations, i.e., as the IoU threshold increases, the AP gap gradually increases

3. Receptive Field Enhancement (RFE)

- Current networks possess square receptive fields

- Mismatch between receptive fields and aspect ratio of faces affect the detection performance

- Propose RFE to diversify receptive fields before predicting classes and locations

- RFE replaces the middle two convolution layers in the class and box subnet of RetinaNet
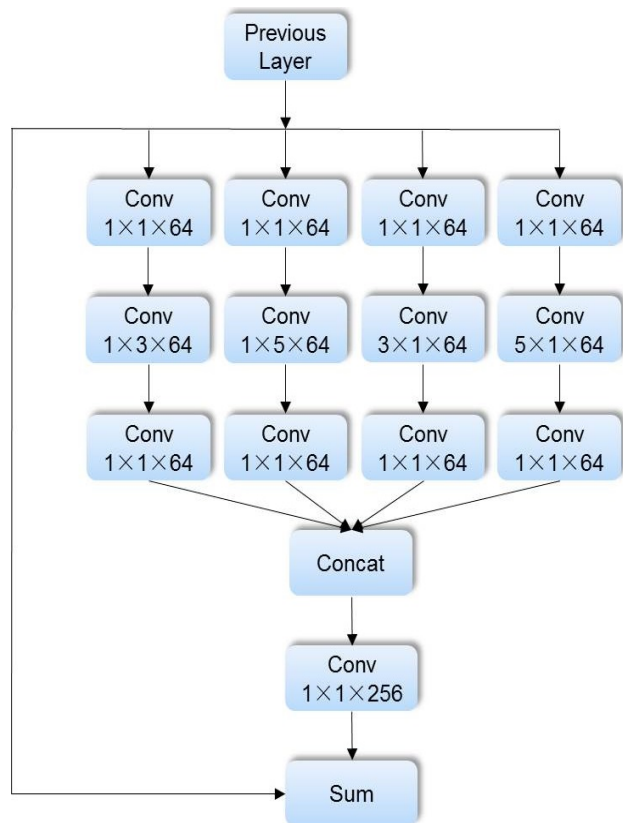
Training Detail

- Backbone: ResNet-50 with 6-level FPN

- Loss: sigmoid focal loss + smooth L1 loss

- Data augmentation: color distortions, random cropping, random flipping

- Anchor design: two specific scales $(2, 2\sqrt{2})$ and one aspect ratio (1.25)

- SGD, 0.9 momentum, 0.0001 weight decay, batch size 32

- Learning rate to 0.01, 0.001 and 0.0001 for the 100, 20 and 10 epochs

Result

- Ablation Study

Component	SRN				
STC		✓		✓	✓
STR			✓	✓	✓
RFE					✓
Easy subset	95.1	95.3	95.9	96.1	**96.4**
Medium subset	93.9	94.4	94.8	95.0	**95.3**
Hard subset	88.0	89.4	88.8	90.1	**90.2**

Result

- Evaluation on Benchmark

Dataset	Criterion	Value				
AFW	Average Precision (AP)	99.87				
PASCAL face	Average Precision (AP)	99.09				
FDDB	True Positive Rate @ False Positive=1000	98.8				
WIDER FACE	Average Precision (AP)	Validation	96.4 (Easy)	95.3 (Medium)	90.2 (Hard)
		Test	95.9 (Easy)	94.9 (Medium)	89.7 (Hard)

Related Paper

Selective Refinement Network for High Performance Face Detection

Cheng Chi[1,3*], Shifeng Zhang[2,3*†], Junliang Xing[2,3], Zhen Lei[2,3], Stan Z. Li[2,3], Xudong Zou[1,3]

[1] Institute of Electronics, Chinese Academy of Sciences, Beijing, China

[2] CBSR & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China

[3] University of Chinese Academy of Sciences, Beijing, China

chicheng15@mails.ucas.ac.cn, {shifeng.zhang,jlxing,zlei,szli}@nlpr.ia.ac.cn, xdzou@mail.ie.ac.cn

- Have been submitted to ARXIV

- Will be announced on September 11

- Refer to it for the details

Thank you!