Selective Refinement Network for High Performance Face Detection

Cheng Chi, Shifeng Zhang, Zhen Lei, Stan Z. Li

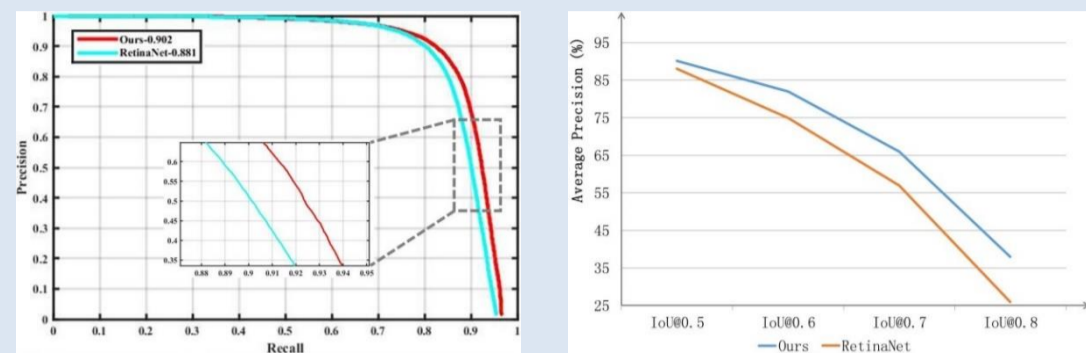University of Chinese Academy of Sciences

Motivation

**(a) Low recall efficiency**   **(b) Low location accuracy**

➢ High performance face detection remains a very challenging problem, especially when there exists many tiny faces.

➢ *Recall efficiency*: number of false positives needs to be reduced at the high recall rates.

➢ *Location accuracy*: accuracy of the bounding box location needs to be improved.

Contribution

➢ Presents a novel single-shot face detector, named **Selective Refinement Network (SRN)**, which introduces novel two-step classification and regression operations selectively to reduce false positives and improve location accuracy simultaneously.

➢ Presents a **Selective Two-step Classification (STC)** module to filter out most simple negative samples to reduce the classification search space.

➢ Designs a **Selective Two-step Regression (STR)** module to provide better initialization for the subsequent regressor.

➢ Introduces a **Receptive Field Enhancement (RFE)** module to provide more diverse receptive fields for detecting extreme-pose faces.

➢ Achieves state-of-the-art results on AFW, PASCAL face, FDDB, and WIDER FACE datasets.

Architecture

Selective Two-step Classification (STC)   Selective Two-step Regression (STR)
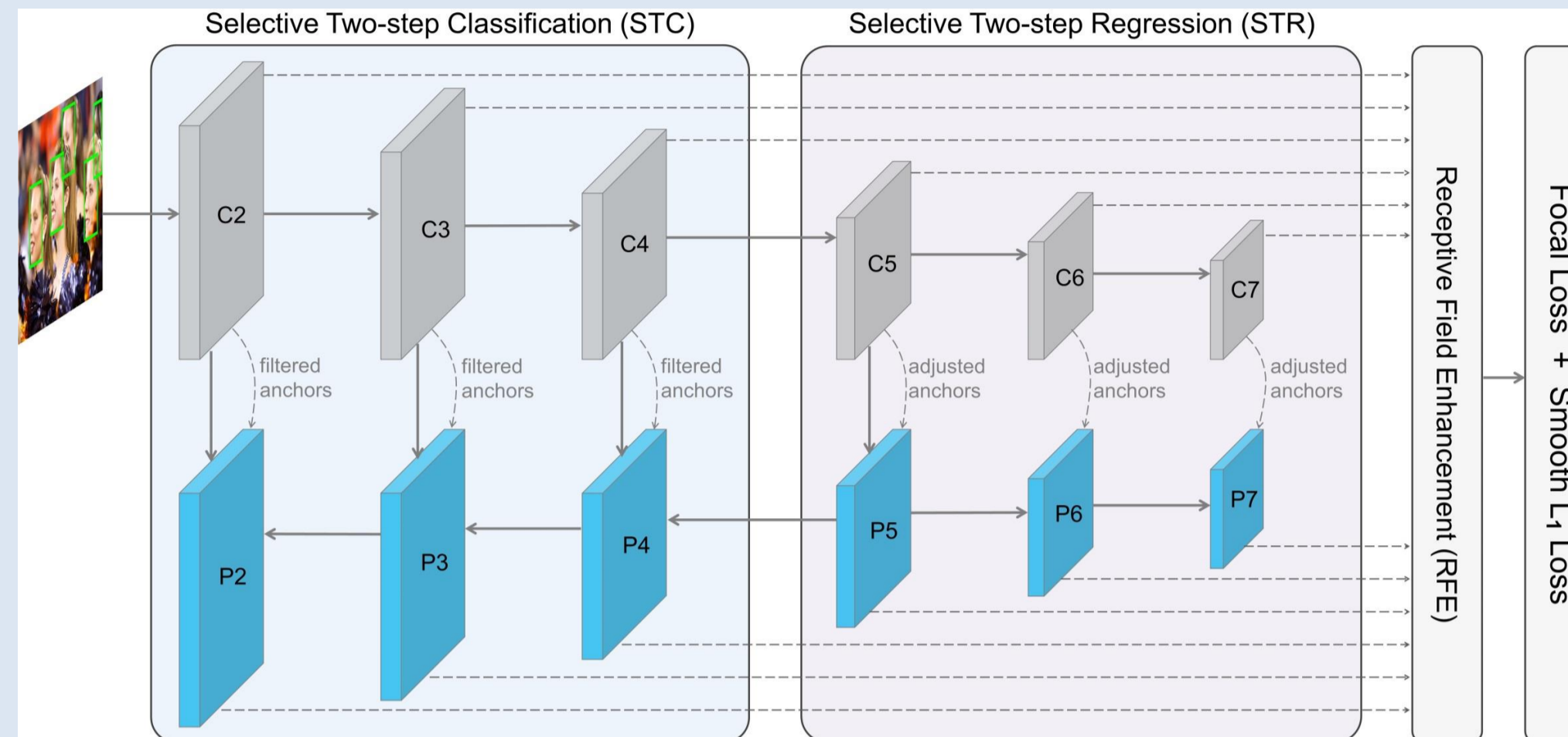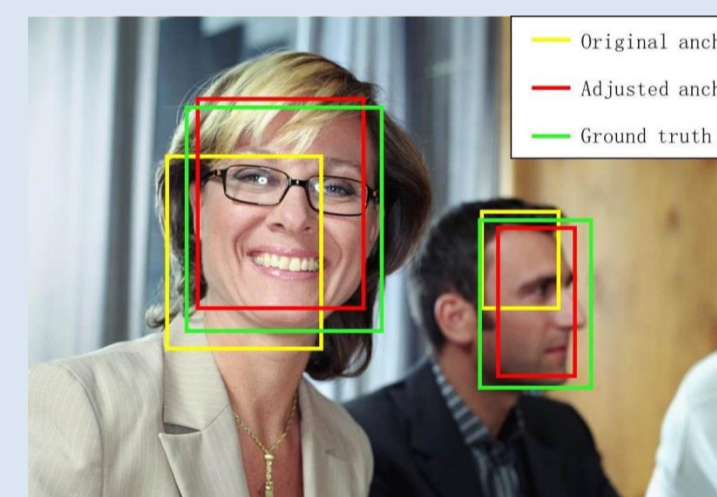
➢ **Selective Two-step Classification (STC)**
  - Selects C2, C3, C4, P2, P3, and P4 to perform two-step classification.
  - To filter out most simple negative samples from low level layers to reduce the classification search space.
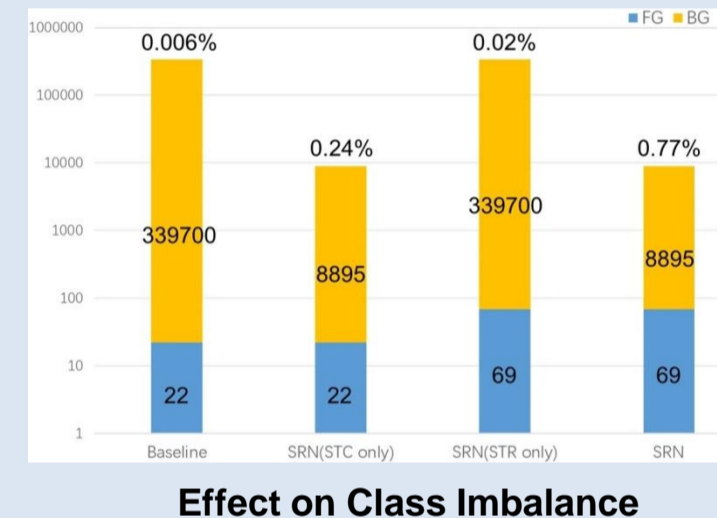
➢ **Selective Two-step Regression (STR)**
  - Selects C5, C6, C7, P5, P6, and P7 to conduct two-step regression.
  - To coarsely adjust the locations and sizes of anchors from high level layers to provide better initialization for the subsequent regressor.

➢ **Receptive Field Enhancement (RFE)**
  - Inspired by the Inception block.
  - By combining multiple branches with different shapes of the kernel size, RFE is able to obtain more diverse receptive fields, which helps to detect faces with large aspect ratios.

Effect on Class Imbalance

Adjusted Anchor
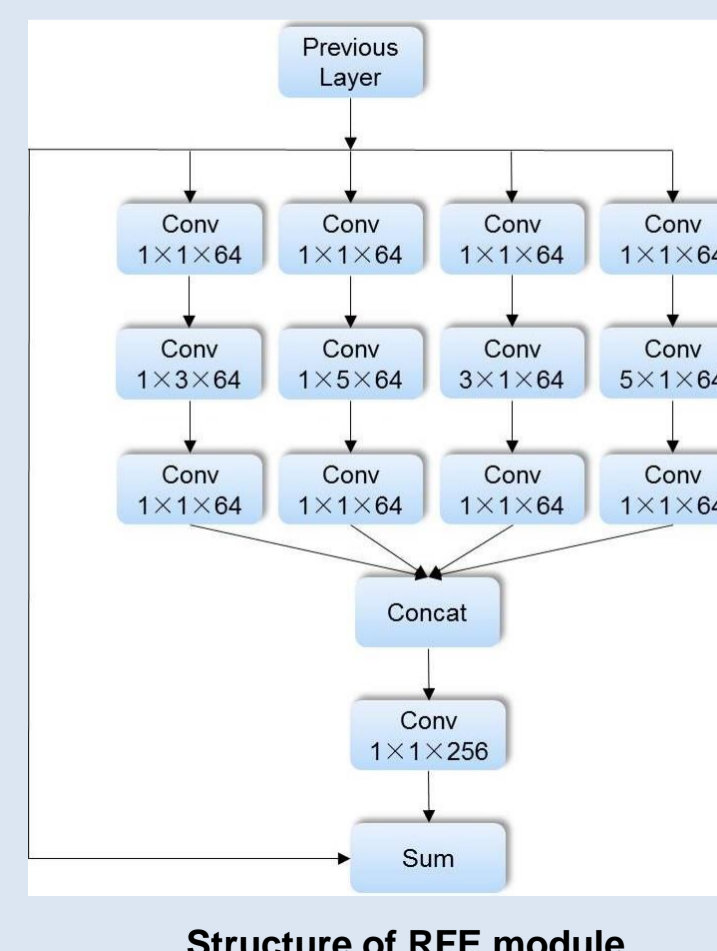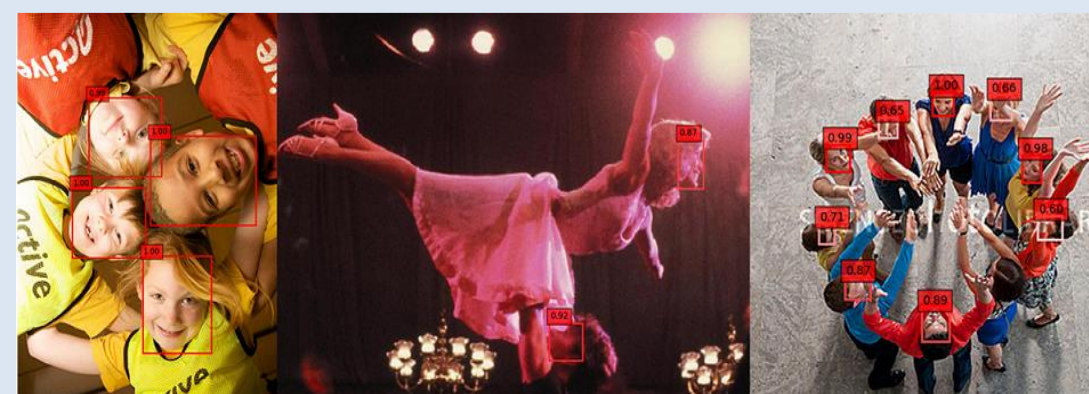
Structure of RFE module

Model Analysis

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**

Effectiveness of various designs on the AP performance

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

AP performance of the two-step classification Applied to each pyramid level

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

AP performance of the two-step regression Applied to each pyramid level

Results on Benchmarks

	Easy	Medium	Hard
WIDER FACE Val	**96.4**	**95.3**	**90.2**
WIDER FACE Test	**95.9**	**94.9**	**89.7**