

# FACEALL-BUPT in ILSVRC2015



Yue WU<sup>1</sup>, Kun HU<sup>1</sup>, Yuxuan LIU<sup>1</sup>, Xuankun HUANG<sup>1</sup>, Jiangqi ZHANG<sup>1</sup>, Hongliang BAI<sup>2</sup>, Wenjian FENG<sup>2</sup>, Tao FU<sup>2</sup>, Yuan DONG<sup>1</sup>  
<sup>1</sup>Beijing University of Posts and Telecommunications, <sup>2</sup>Beijing Faceall co., LTD

## Object DET

- Selective Search + Fast RCNN
- AlexNet & GoogLeNet
- Pre-trained CNN models in training set of CLS
- Fine-tuned in training set of DET
- Different position of the ROI pooling layer
- Large input size & ROI pooled size
  - input size: 768x1280 from 600x1000
  - Pooled size: 4x4 from 1x1
- Three models ensemble

## Object CLS-LOC

- Selective Search + GoogLeNet
- Classification
  - GoogLeNet with BN
  - Multi-scale training & testing
  - Three models ensemble
- Localization
  - Pre-trained in image-level annotations
  - Fine-tuned in object-level annotations
  - Bounding box refinement by merging overlapped regions
  - Two models ensemble

## Scene CLS

- Small GoogLeNet
- Input size: 128x128
- Larger CONV stride
- Reduced kernel numbers
- Trained 13 epochs

## Object DET from Video

- Selective Search + Fast RCNN + TLD
- Key frames selection for object detection
  - train: 52,922 of 1,322,686 frames
  - val : 6,861 of 176,126 frames
  - test : 12,329 of 315,176 frames
- Object tracking
  - Tracking-Learning-Detection(TLD)
  - Forward-Backward tracking from a detected object
- Two models ensemble

Entry ID	MAP on VAL/TEST
1	34.9%/34.8%
2	34.0%/34.3%
3	37.8%/37.9%
4	38.7%/38.6%

Entry ID	TOP-5 CLS ERROR ON VAL/TEST	TOP-5 LOC ERROR ON VAL/TEST
1	7.03%	32.9%/32.8%
2	7.03%	32.8%/32.8%
3	7.03%	31.5%/31.5%

Entry ID	TOP-5 ERROR ON VAL/TEST
1	37.2%/38.9%
2	39.0%/40.7%
3	31.4%/32.7%

Entry ID	MAP ON VAL/TEST
1	19.1%/16.2%
2	24.6%/21.8%
3	25.3%/22.2%

### References

- [1] Going Deeper With Convolutions, CVPR 2014
- [2] Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv, 2015
- [3] Very deep convolutional networks for large-scale image recognition, ICLR 2015
- [4] Fast R-CNN, ICCV 2015
- [5] Tracking-Learning-Detection, TPAMI 2012