

ScratchDet: Training Single-Shot Object Detectors from Scratch

Rui Zhu^{1,4*}, Shifeng Zhang^{2*}, Xiaobo Wang¹, Longyin Wen³, Hailin Shi¹, Liefeng Bo³, Tao Mei¹

¹JD AI Research, ²CASIA UCAS, ³JD Digits, ⁴Sun Yat-sen University

Section 1: Motivation

➤ Previous fine-tuned strategies:

- High computational cost of training on ImageNet
- Learning bias from classification to detection
- Inconvenient to change the architecture of networks

➤ DSOD is the first attempt to train detectors from scratch:

- Limited by the DenseNet architecture
- Fail to converge for two-stage detectors
- Performance gap compared with fine-tuned ones.

➤ Targets of training from scratch:

- Free the architecture limitations
- Guarantee the training convergence
- Performance as good as fine-tuned strategies

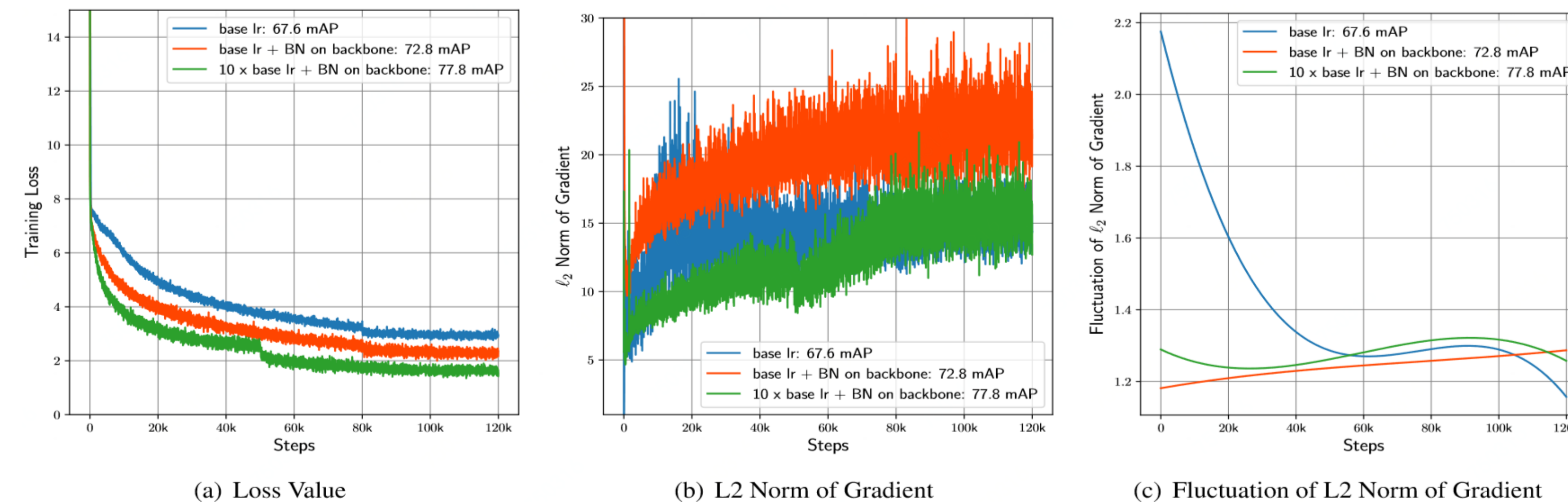
Section 2: Contribution

➤ We deeply analyze the effect of **Batch Normalization** for train-from-scratch.

➤ We design a new **root block** to keep the abundant information for small object detection.

➤ We conduct **extensive experiments** on several benchmarks to validate the effectiveness of our method.

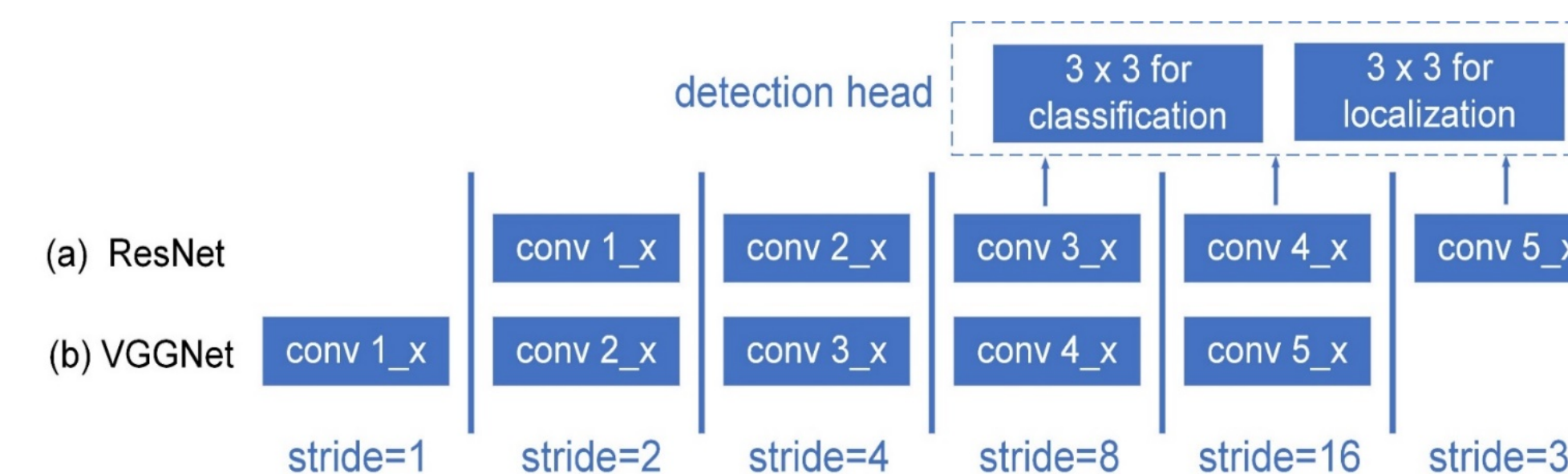
Section 3: ScratchDet



➤ BatchNorm for train-from-scratch:

- Smoother optimization landscape
- More stable gradients
- Enabling larger learning rate

With BatchNorm, we can modify the network structure without the pre-trained restrictions !!!

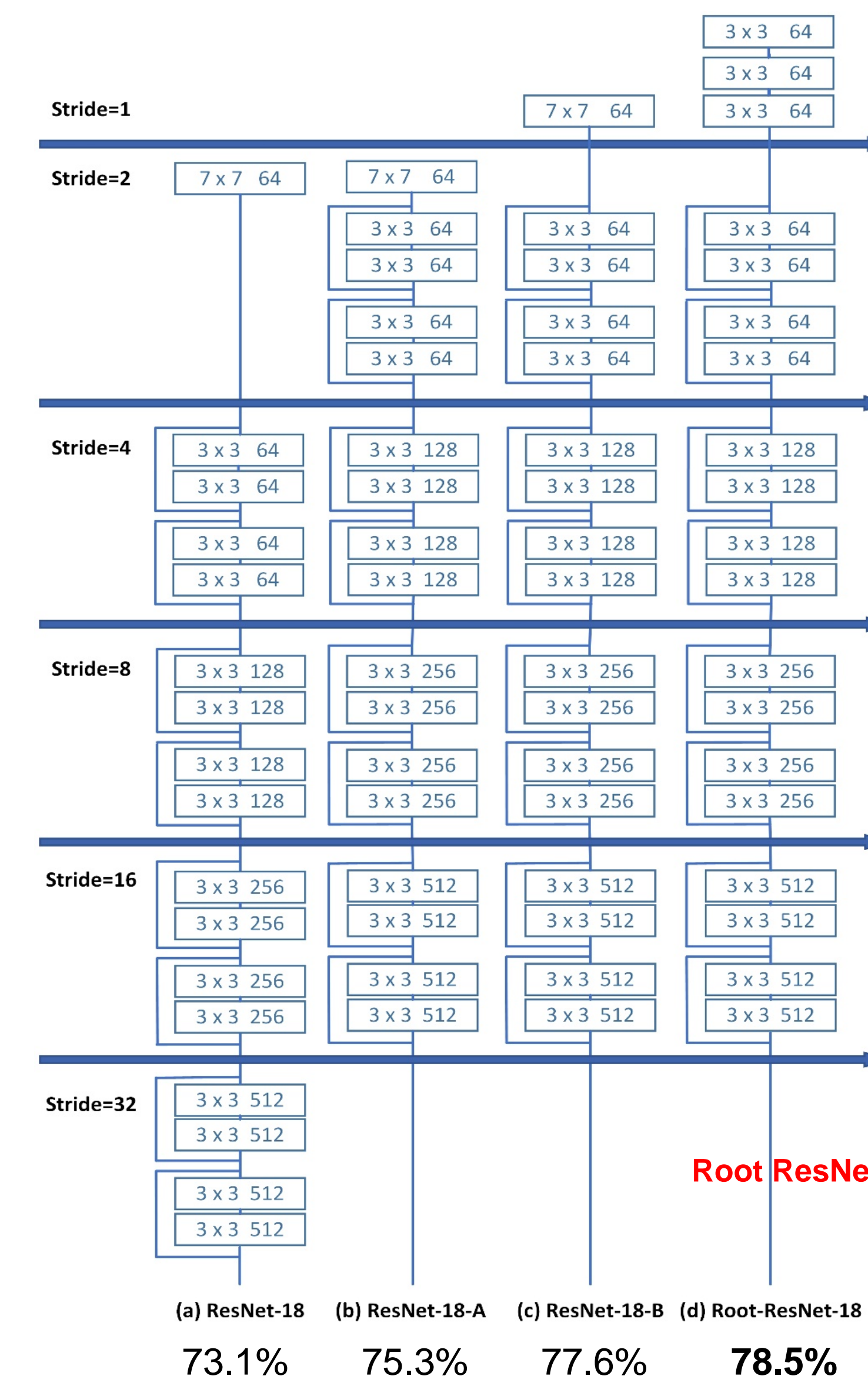


➤ Remove down-sampling operation in the first convolutional layer (i.e., stride=1)

Root Block !!!

➤ Replace the 7×7 convolutional filter by a stack of 3×3 convolutional filters

With Root-Block, we can achieve the performance as good as fine-tuned detectors !!!



Section 4: Experiments

Effectiveness of BatchNorm

Component	SSD300-from-scratch				
50x large learning rate?	✓				
10x large learning rate?		✓			
BN on backbone?	✓	✓	✓		
BN on detection head?	✓	✓	✓	✓	
mAP (%) on VOC07 test	78.7	77.3	71.8	71.0	67.6

Effectiveness of Root-Block

First conv layer	Root block	mAP
with downsampling	1#: 7x7	73.1
	3#: 3x3	75.4
without downsampling	1#: 7x7	77.6
	3#: 3x3	78.5

Performance on Benchmarks

Method	VOC07 test		VOC12 test		COCO test-dev		
	0712	0712+COCO	0712	0712+COCO	Trainval35k		
	0.5	0.5	0.5	0.5	0.5:0.95	0.5	0.75
Single-scale	80.4	84.0	78.5	82.1	32.7	52.0	34.9
Multi-scale	84.1	86.3	83.6	86.3	39.1	59.2	42.6

Codes and models: <https://github.com/KimSoybean/ScratchDet>