

Solution for Learning from Large-Scale Web Images

PCI Team
June 15, 2020



CONTENTS

1

Datasets & Challenges

2

Solution Techniques

3

Experiment Results

4

Our Team

Datasets & Challenges

Training set:

Google 8,366,429

Flickr 7,710,236

16,076,665



Validation set:
294,099

Testing set:
294,099



Imbalanced

There are 53,647 images in "human" class but merely 98 images belong to the "driver" class.

Huge

16M images & 5,000 classes

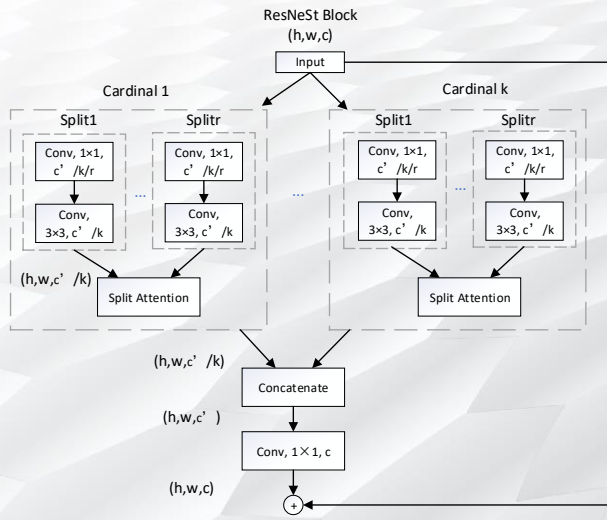
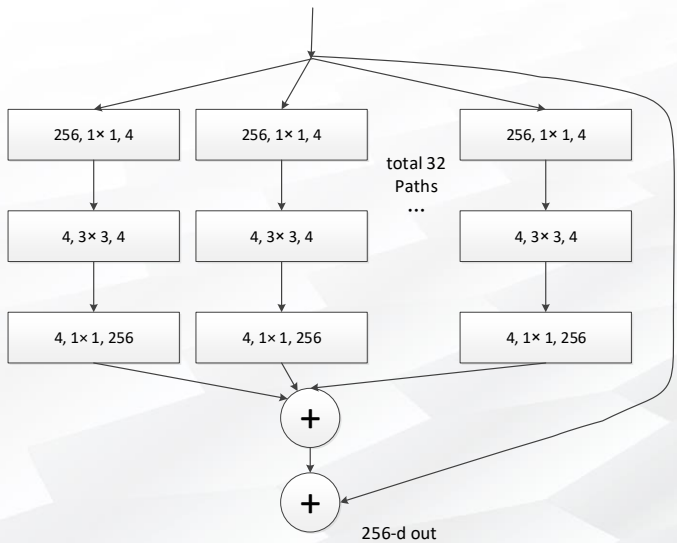
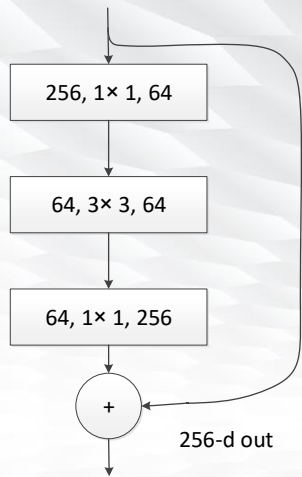
Noisy

Un-labeled training images

Solution Techniques 1/4

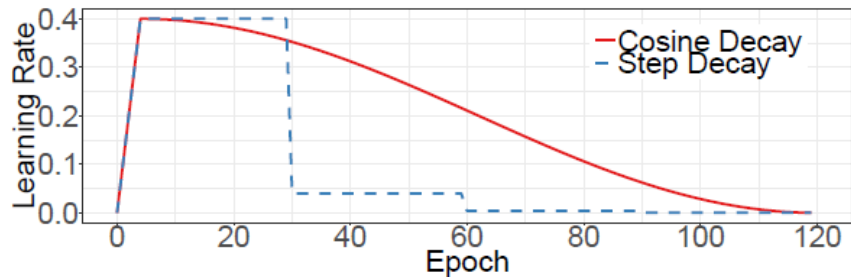
Model architecture

- ResNet101, ResNet152
- ResNext101
- ResNest101



□ Training strategies

- Large batch training
- Learning rate schedule: warmup+cosine
- Data augment: mixup, label smooth
- Low precision training: fp16



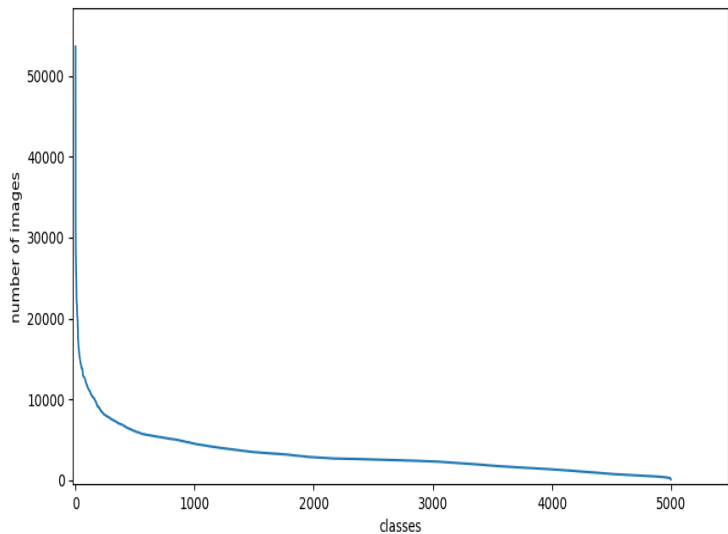
(a) Learning Rate Schedule



Solution Techniques 3/4

Techniques for the imbalanced data:

- Reweighting class
- Adding more weights to the long-tail classes



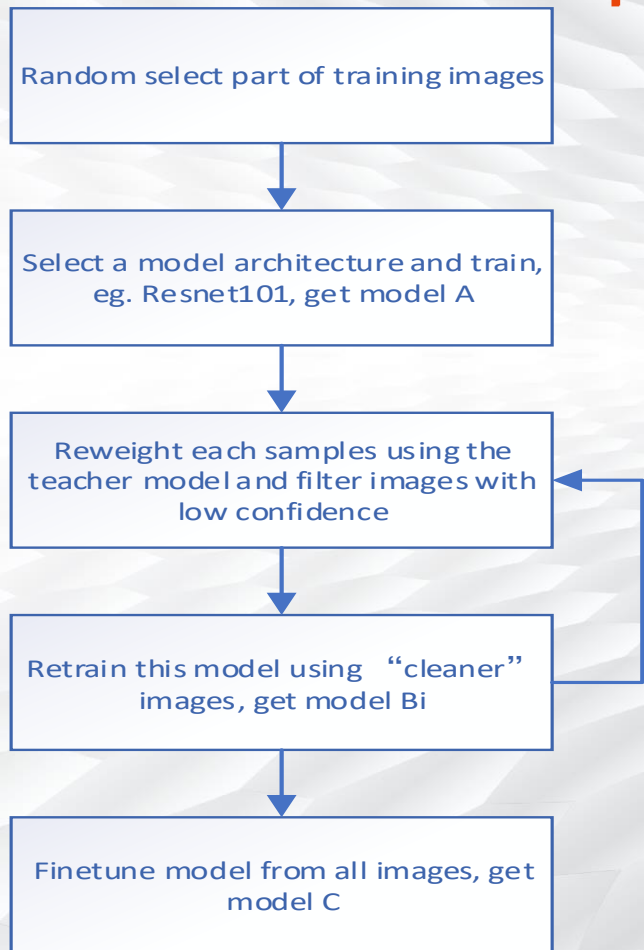
Techniques for the noisy data:

- Reweighting images
- Using trained model to generate the weight for each image



Solution Techniques 4/4

- Solution alternative process
 - Train from the auto “cleaner” data
- Testing tricks
 - Multi-model ensemble
 - Multi-scale test
 - Multi-crop test



Experiment Results 1/2

ResNet 101

ResNet 101	Input	Top1-val	Top5-val
A	224x224	51.78%	74.76%
B1	224x224	55.52%	77.78% (+3.02%)
B2	224x224	55.87%	78.13% (+0.35%)
C	224x224	54.74%	78.69% (+0.56%)

ResNet 152

ResNet 152	Input	Top1-val	Top5-val
B1	224x224	55.84%	78.01% (+3.25%)
B2	224x224	56.16%	78.36% (+0.35%)
C	224x224	56.77%	78.83% (+0.47%)

Experiment Results 2/2

ResNext 101

ResNext 101	Input	Top1-val	Top5-val
B1	224x224	55.63%	77.91% (+3.15%)
C	224x224	56.58%	78.63% (+0.72%)

ResNest 101

ResNest 101	Input	Top1-val	Top5-val
B1	224x224	56.36%	78.53% (+3.77%)
C	224x224	57.23%	79.23% (+0.70%)

- Using multi-model ensemble, multi-scale, and multi-crop tests, final top5-test at 79.88%

PCITECH Global Intelligent Technology Research Institute

Group 1: Perception intelligence

Visual image algorithm development, algorithm chip design and visual AI platform development.

Group 2: Cognitive intelligence

Big data, knowledge graph, and industrial internet platforms.

Group 3: Intelligent transportation systems

New technologies in urban transportation systems, and future mobility solutions for mega cities.



Group 4: Urban rail transit

New technologies for urban rail transit systems.

Group 5: 5G technology applications

Industrial solutions and support of mobile communications, promote industrial 5G applications.

Group 6: Special projects

Advanced technologies and products for solving hard-core problems in emerging areas.

THANK YOU

PCI 佳都科技

城市慧变得更好!
perfect city

