Deep learning and realistic datasets

Zhongqi Miao, Ziwei Liu
Summary

- Background
- Large-scale long-tailed recognition in an open world
- Open compound domain adaptation
Background

- Deep learning looks so powerful!!
Problem

• Even the state-of-the-art methods are not good enough to handle realistic data in realistic settings!
Problem

Training

Train $T_0$

CNN $T_0$

Classifier $T_0$

$T_0$
Problem

Evaluation

Validate $T_0$ → CNN $T_0$ → Classifier $T_0$
Problem

Inference

New Data $T_1$ → CNN $T_0$ → Classifier $T_0$
Problem

Inference

New Data

Classifier

CNN

\( T_0 \)

\( T_1 \)
Problems

• Long-tailed
• Open-ended
• Multi-domain
Problem

Fine-tuning
Problem

Fine-tuning requires large amount of labeled data.

Fine-tuning
Problem

Fine-tuning
Large amount of labeled data is even less possible.

Fine-tuning
Long-tailed distribution

- Red-winged Blackbird
- Spotted Sandpiper
- Gun Shots
Long-tailed distribution
Open!

# detections

Categories

Red-winged Blackbird

Spotted Sandpiper

Gun Shots

1000

5

New classes will usually be on this side
Problems

• Long-tailed

• Open-ended

• Multi-domain
Open long-tailed recognition

Ziwei Liu$^{1,2}$, Zhongqi Miao$^2$, Xiaohang Zhan$^1$, Jiayun Wang$^2$, Boqing Gong$^{3,2}$, Stella X. Yu$^2$

1. The Chinese University of Hong Kong
2. UC Berkeley / ICSI
3. Google Inc.

CVPR, 2019, oral
Long-tailed distribution

- Modern deep learning techniques are based on large-scale balanced training datasets:
Long-tailed distribution

Categories

Red-winged Blackbird

Spotted Sandpiper

Gun Shots
Faces [Zhang et al. 2017]

Species [Van Horn et al. 2019]

Places [Wang et al. 2017]

Actions [Zhang et al. 2019]
Open Long-Tailed Recognition

Head Classes

Tail Classes

Open Classes

Open World
Open Long-Tailed Recognition

Imbalanced Classification

Few-shot Learning

Open Set Recognition

Open World

Head Classes

Tail Classes

Open Classes
Imbalanced Classification  
(metric learning, re-sampling, re-weighting)

Few-Shot Learning  
(meta learning, classifier dynamics)

Open Set Recognition  
(distribution rectification, out-of-distribution detection)

Open Long-Tailed Recognition  
(dynamic meta-embedding)

Sensitivity to Novelty  
Avoid Forgetting  
Knowledge Transfer  

Sensitivity to Novelty  
Avoid Forgetting  
Knowledge Transfer
Open Long-Tailed Recognition

- Avoid Forgetting
- Knowledge Transfer

Head Classes  Tail Classes  Open Classes

Open World

Sensitivity to Novelty
Head Classes

Tail Classes

Open Classes

Avoid Forgetting

Knowledge Transfer

Sensitivity to Novelty

bottom-up attention

visual memory

top-down attention

familiarity
(b.1) Input Image  
(b.2) Feature Map of Plain ResNet Model  
(b.3) Feature Map of Our Model  
(b.4) Modulated Attention
Figure 1. Embedding of Plain ResNet Model vs. Embedding of Dynamic Meta-Embedding.
Head Classes

Tail Classes

Open Classes

familiarity

visual memory

bottom-up attention

embedding

rescaled embedding

Tail Class 'African Grey'

Tail Class 'Buckeye'

Open Sample
Modulated Attention ($f^{att}$)

Input Image ($x$)

DIRECT FEATURE

Con. Sel. ($e$)

fc + tanh

Hallucination ($T_{hal}$)

fc + softmax

MEMORY FEATURE

Reachability ($\gamma$)

Min. dis.

OLTR FEATURE

Classifier ($\emptyset$)

cosine norm.

Logits

Memory

Input Image

Modulated Attention

Con. Sel.

Hallucination

Memory Feature

OLTR Feature
ImageNet-LT Benchmark
Absolute Performance Gain: ~20%

Places-LT Benchmark
Absolute Performance Gain: ~10%

MS1M-LT Benchmark
Absolute Performance Gain: ~2%
Few shot
Open compound domain adaptation

Ziwei Liu*, Zhongqi Miao*, Xingang Pan, Xiaohang Zhan, Dahua Lin,
Stella X. Yu, Boqing Gong
The Chinese University of Hong Kong & UC Berkeley / ICSI &
Google Inc.

CVPR, 2020, oral
Simulation

Open World Driving Conditions
Compound Heterogeneous Domains

- Cloudy
- Rainy
- Snowy

Simulation

Source

Target
(a) Unsupervised Domain Adaptation

(b) Multi-Target Domain Adaptation
Class "1"  
Class "7"  
Class "4"  
Class "9"  
Class "0"  
Class "2"  
Class "3"  
Class "8"  
Class "6"  
Class "5"  

Domain "MNIST" (compound)  
Domain "SVHN" (source)  
Domain "SymNum" (open)  
Domain "MNIST-M" (compound)  
Domain "USPS" (open)  
Domain "SWIT" (open)
Curriculum according to Domain Characteristics

Source

Compound Targets

instance-wise curriculum

Open Targets

domain memory

Continuous Adaptation

Domain Disentanglement

Adaptive Knowledge Transfer
Memory-Augmented Domain Indicator

\[ \nu_{\text{transfer}} = \nu_{\text{direct}} + e_{\text{domain}} \otimes \nu_{\text{enhance}} \]
Continuous Adaptation

Source

Simulation

Compound Targets

Open World Driving Conditions

Cloudy

Rainy

Overcast

Open Targets

instance-wise curriculum

domain memory

Domain Disentanglement

Adaptive Knowledge Transfer
C-Digits Benchmark
Absolute Performance Gain: ~5%

C-Faces Benchmark
Absolute Performance Gain: ~10%

C-Driving Benchmark
Absolute Performance Gain: ~2%

C-Mazes Benchmark
Absolute Performance Gain: ~30%
Problem continues

Deep learning system
Problem continues

Realistic machine learning system
Thank you!