Deep learning and realistic datasets

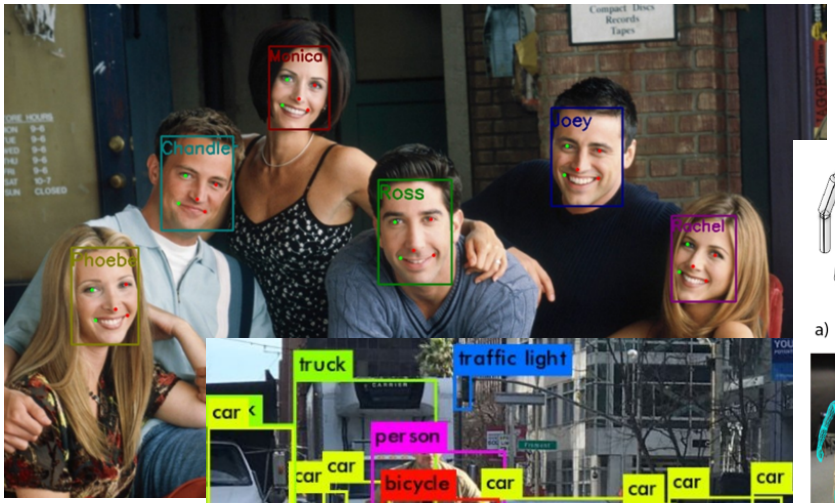
Zhongqi Miao, Ziwei Liu

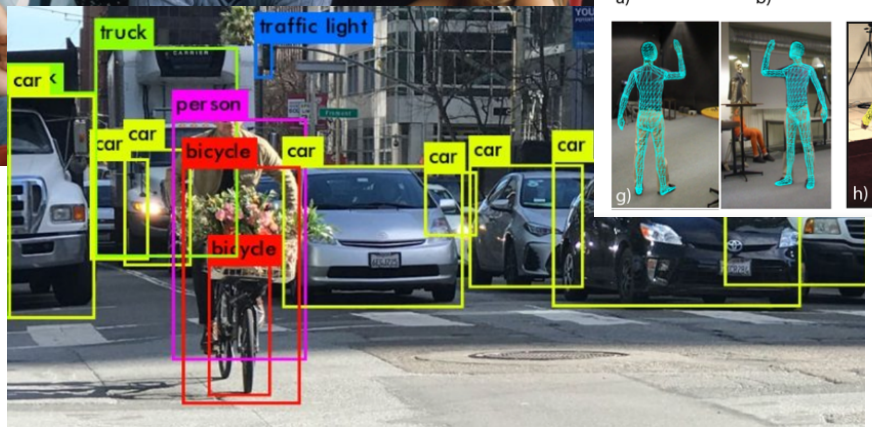
Summary

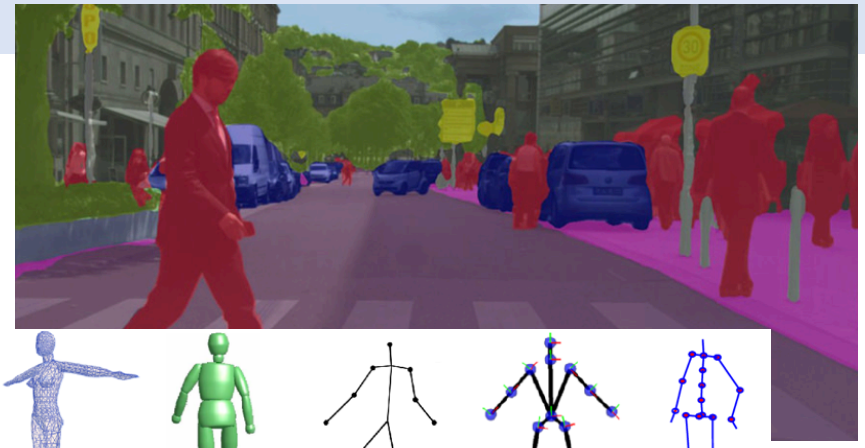
- Background
- Large-scale long-tailed recognition in an open world
- Open compound domain adaptation

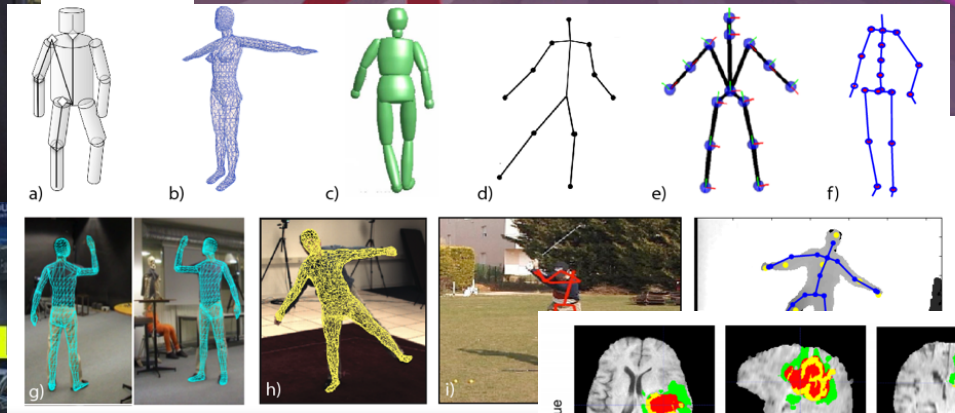
Background

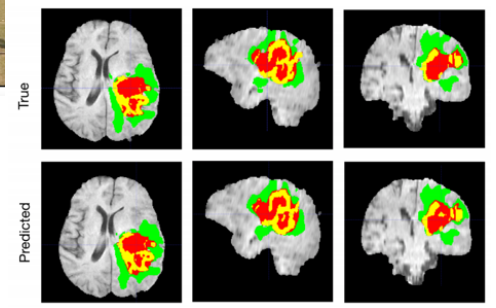
- Deep learning looks so powerful!!









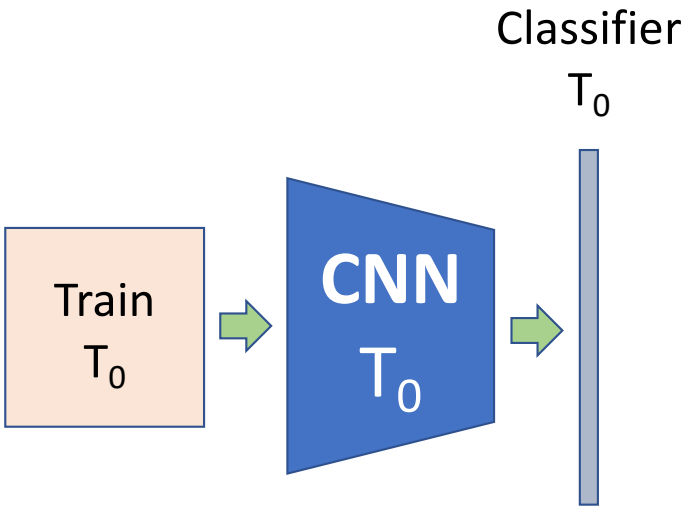


Problem

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

Problem

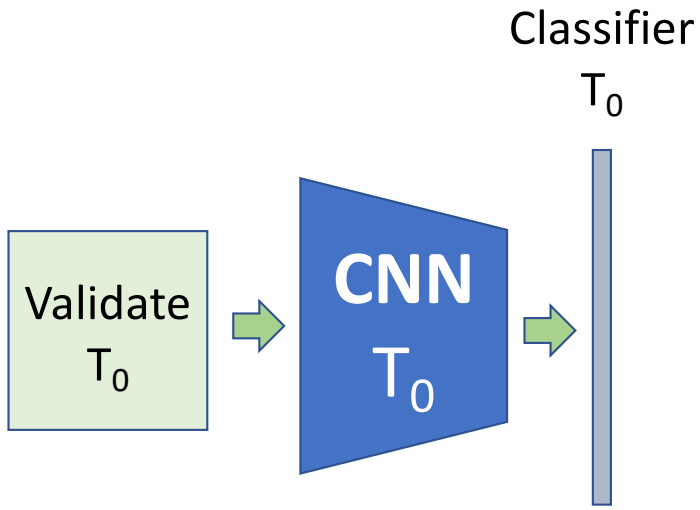
Training



T_0

Problem

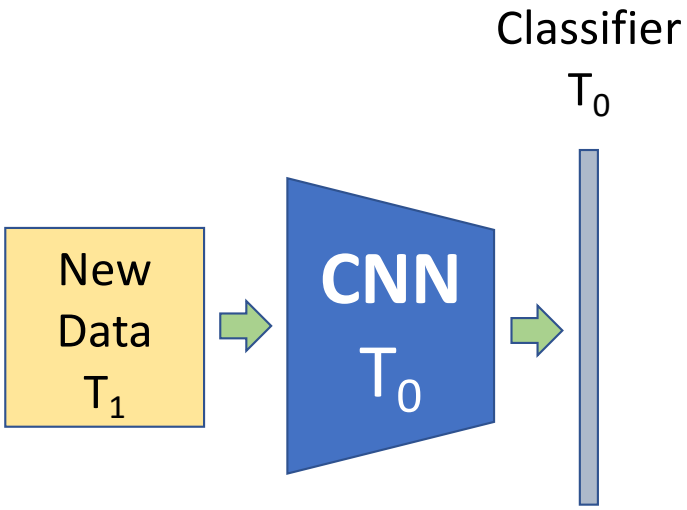
Evaluation



T_0

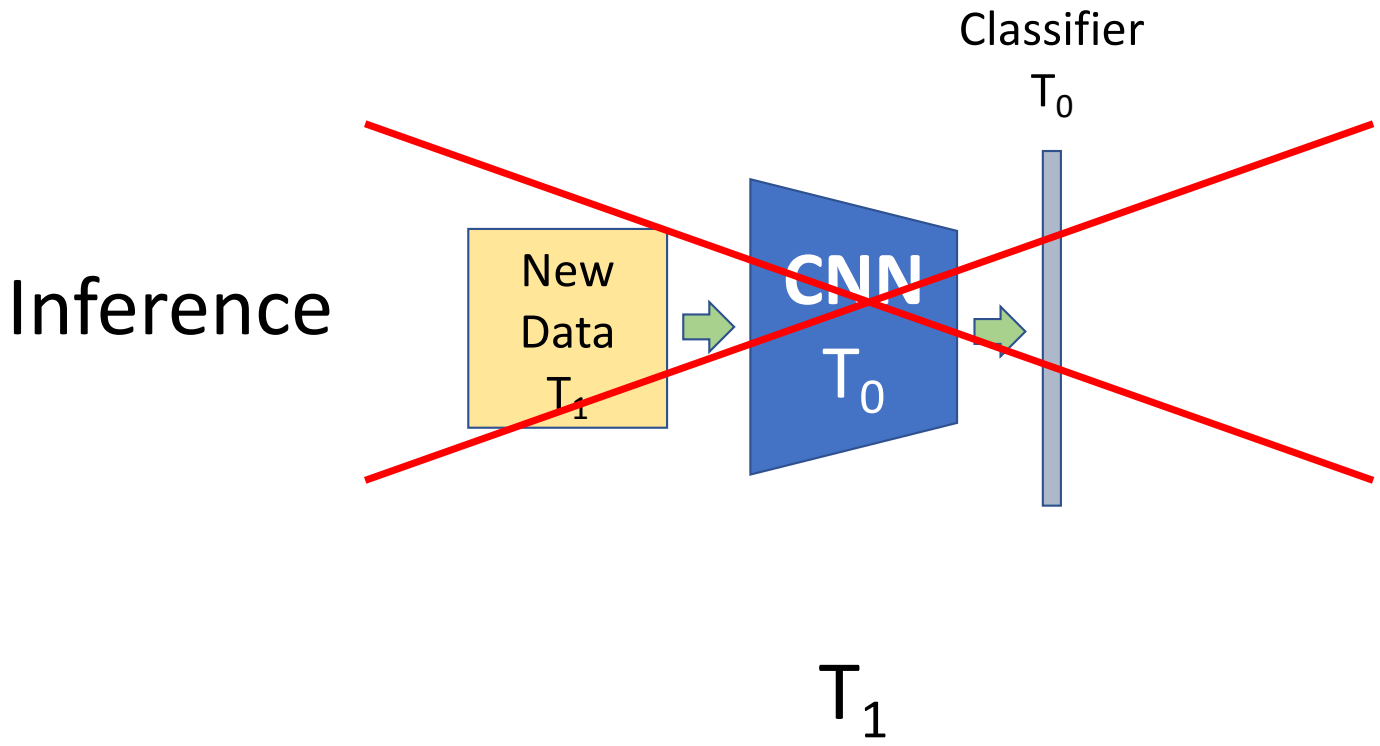
Problem

Inference



T_1

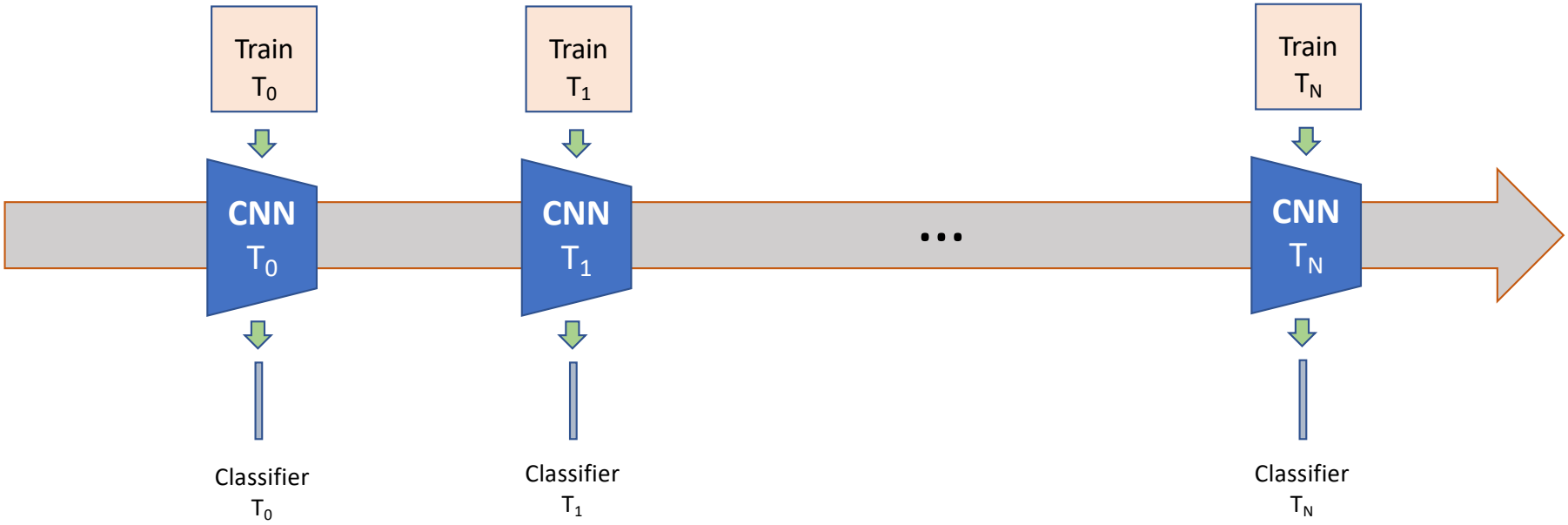
Problem



Problems

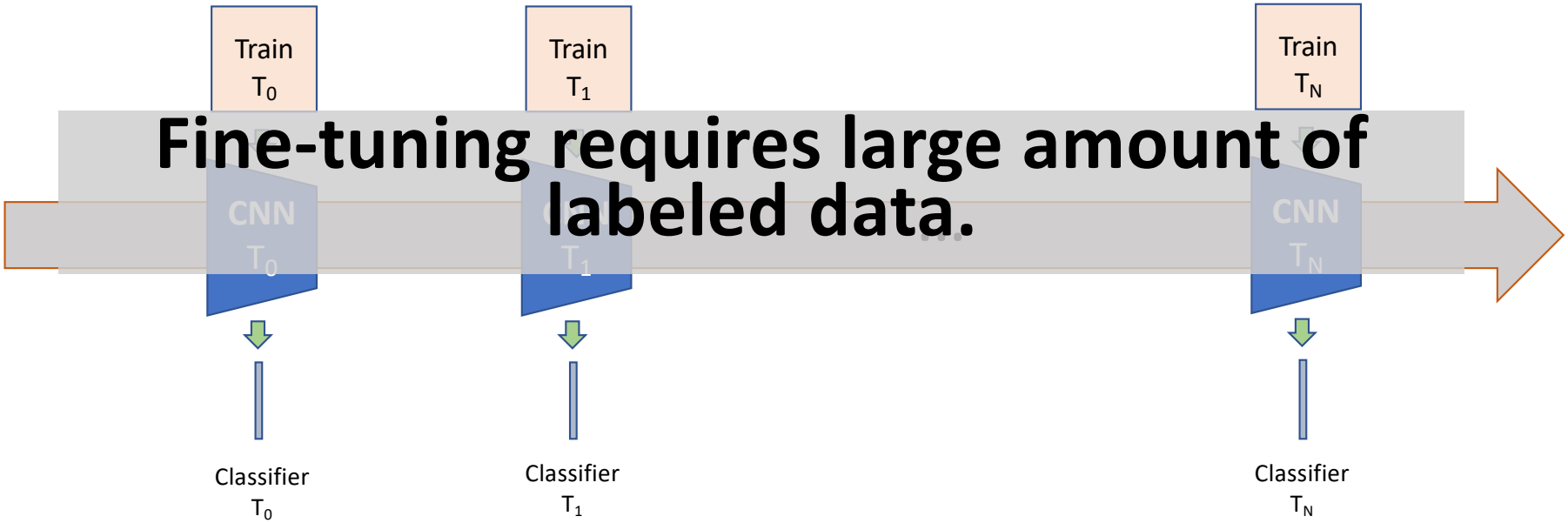
- Long-tailed
- Open-ended
- Multi-domain

Problem



Fine-tuning

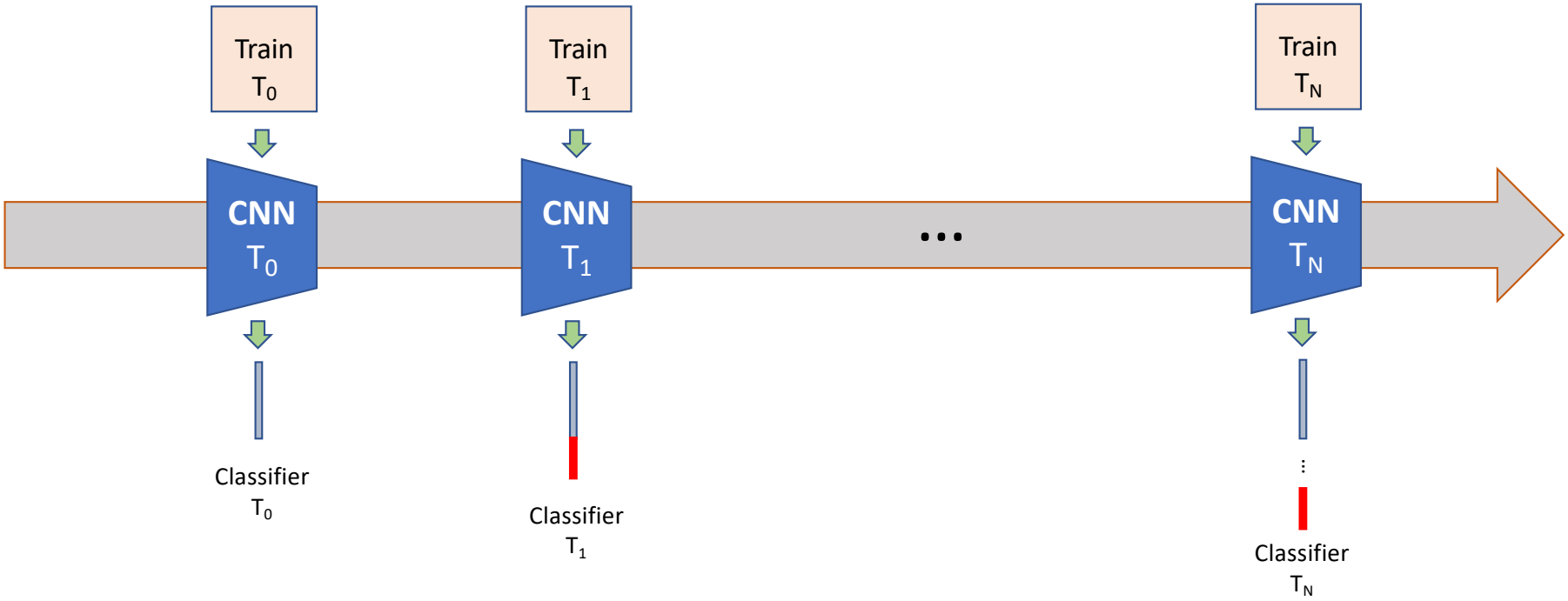
Problem



Fine-tuning requires large amount of labeled data.

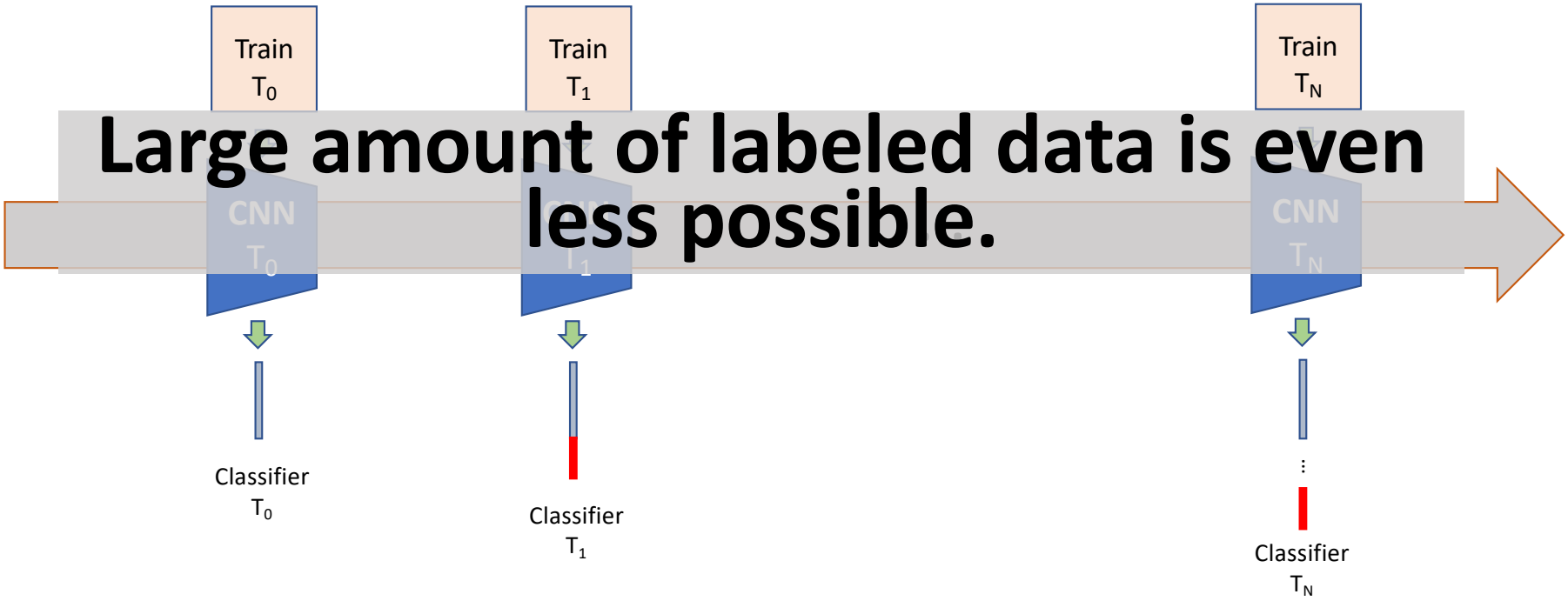
Fine-tuning

Problem



Fine-tuning

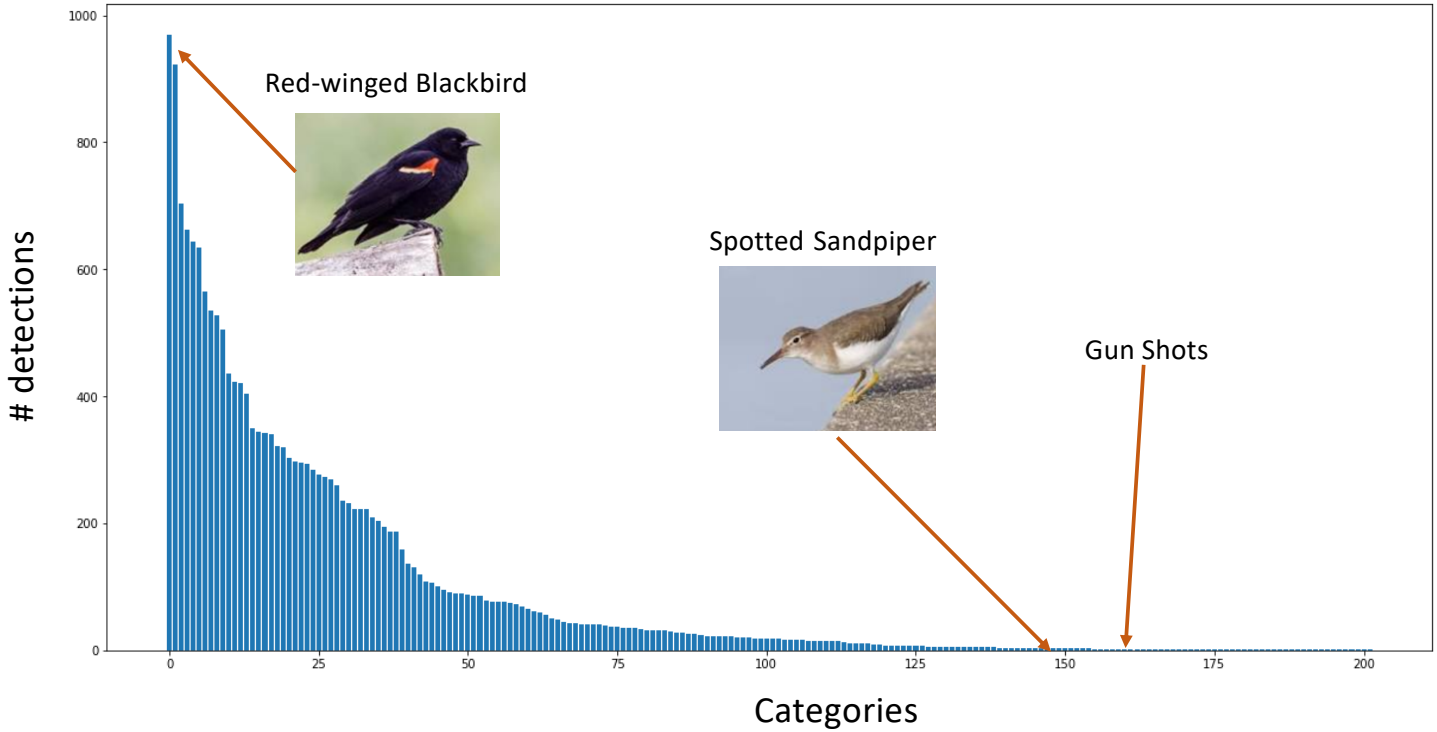
Problem



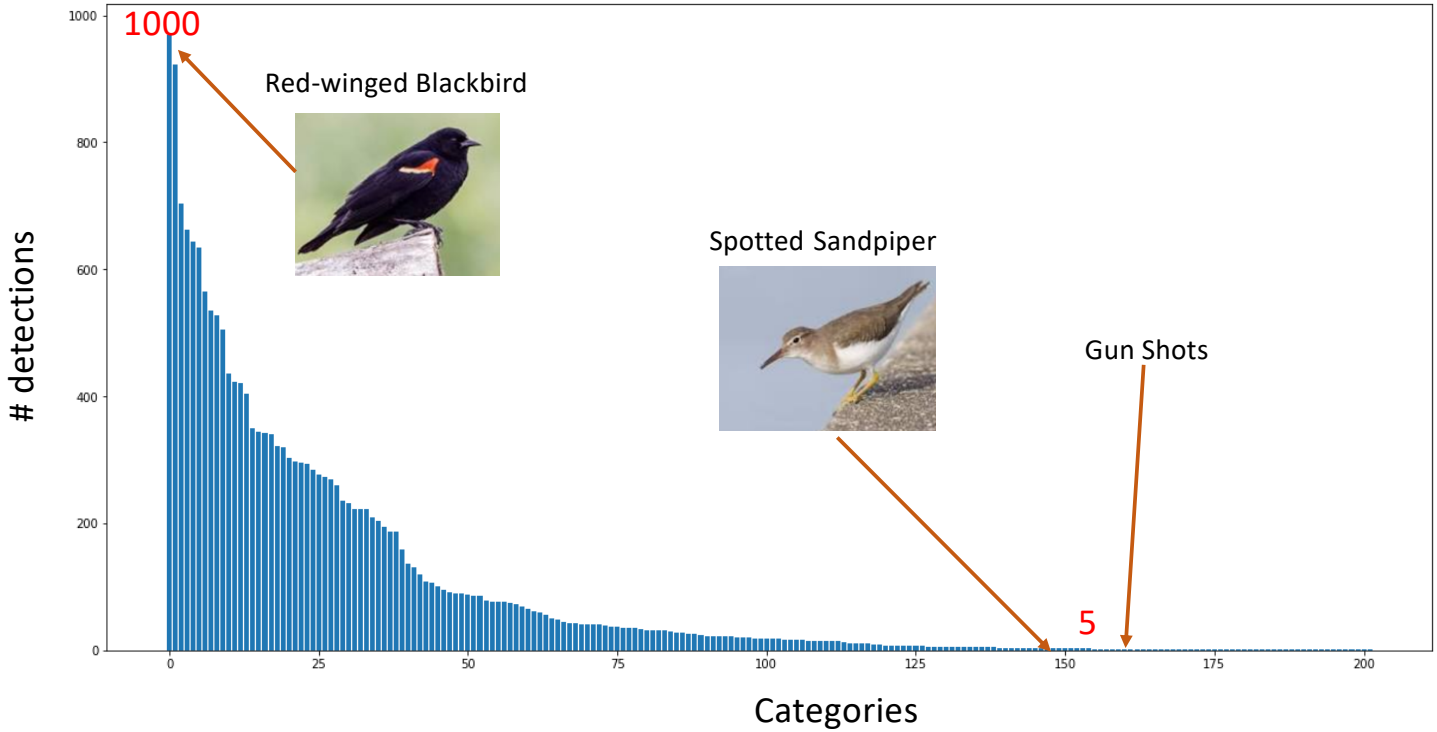
Large amount of labeled data is even less possible.

Fine-tuning

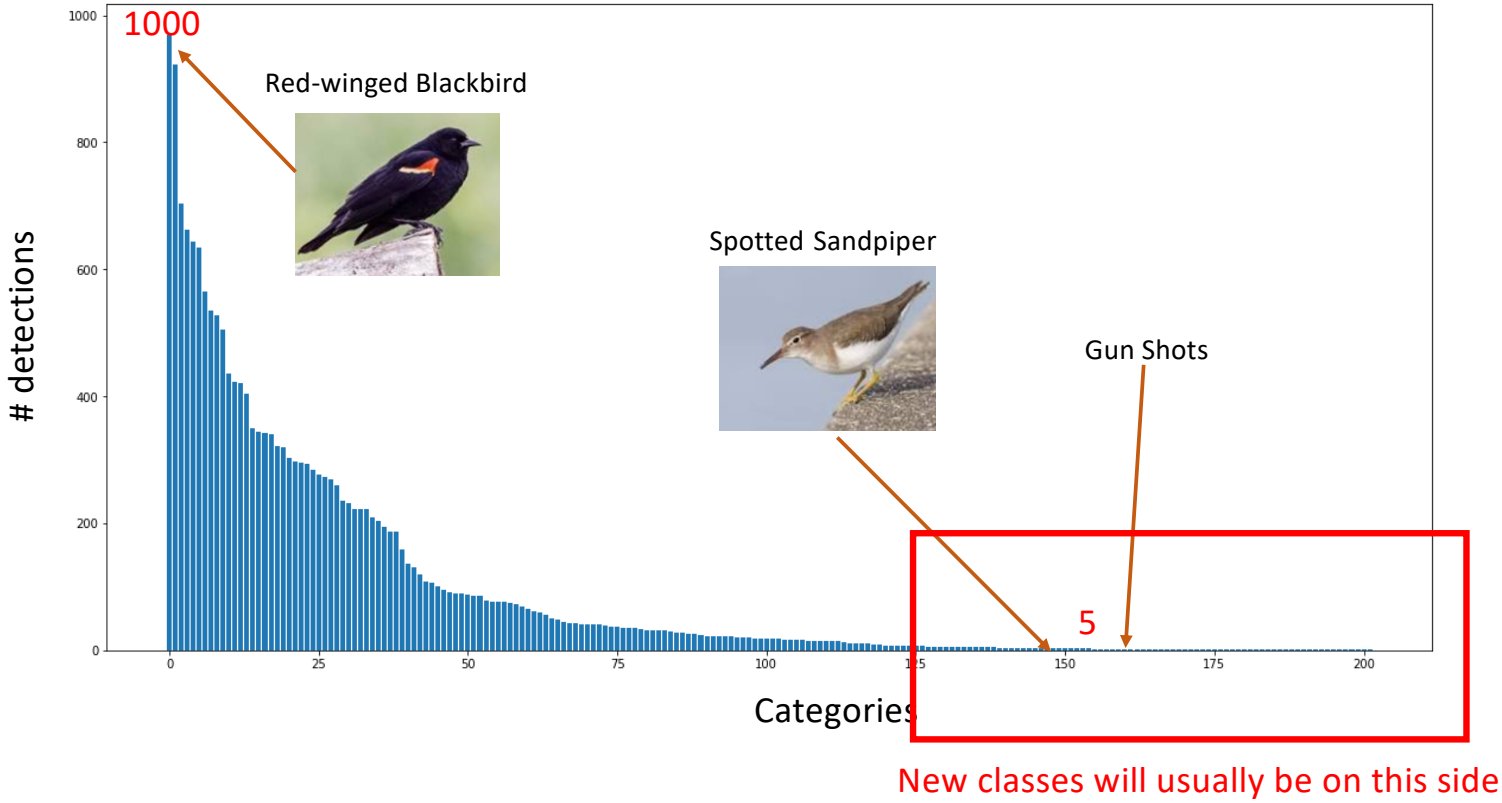
Long-tailed distribution



Long-tailed distribution



Open!


1000
Red-winged Blackbird
Spotted Sandpiper
Gun Shots
5
detections
Categories
0
200
400
600
800
1000
0
25
50
75
100
125
150
175
200
New classes will usually be on this side

Problems

- Long-tailed
- Open-ended
- Multi-domain

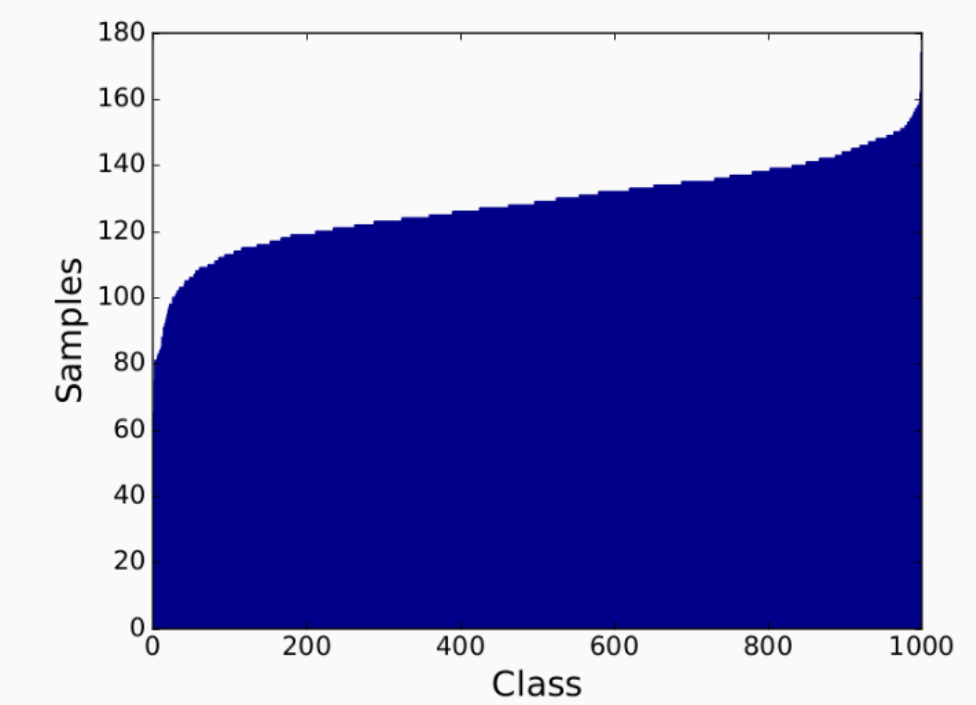
Open long-tailed recognition

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.

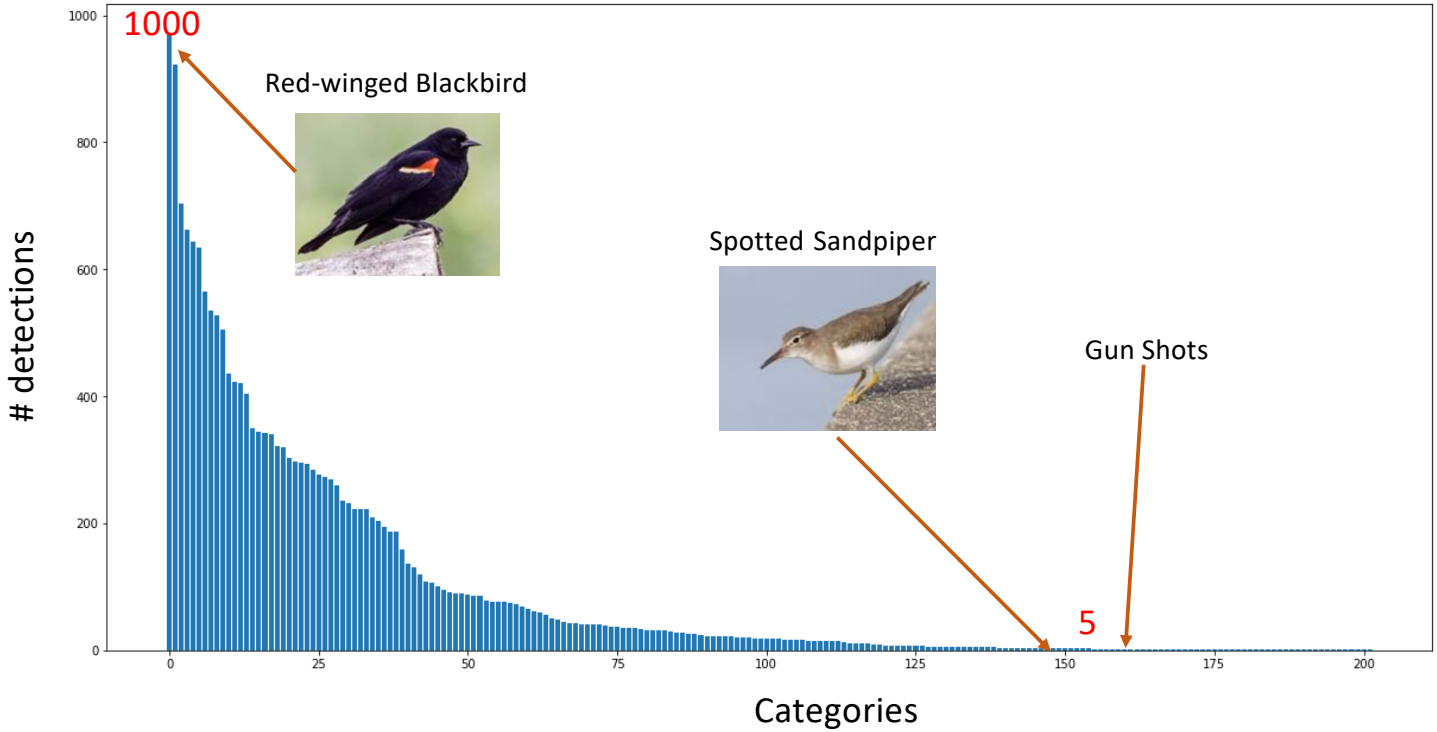
CVPR, 2019, oral

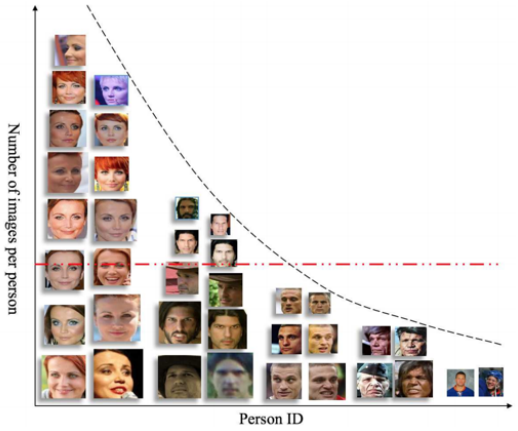
Long-tailed distribution

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

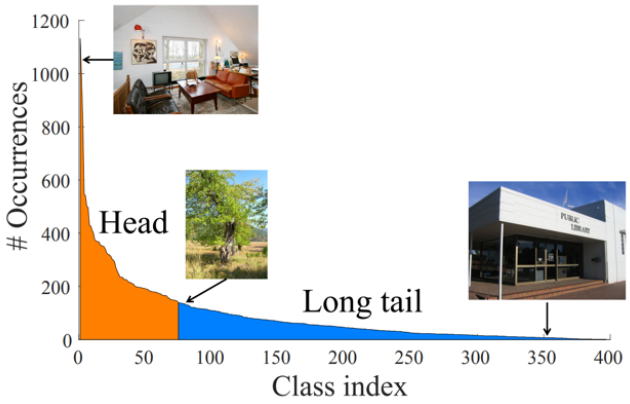


Long-tailed distribution

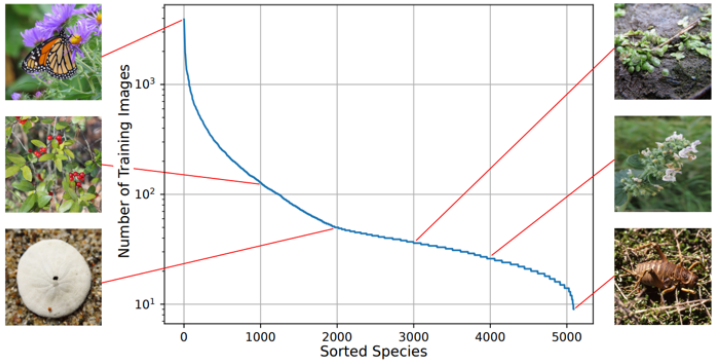




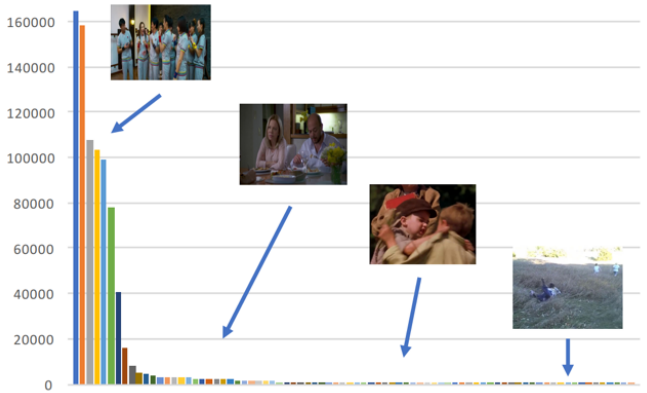
Faces [Zhang et al. 2017]



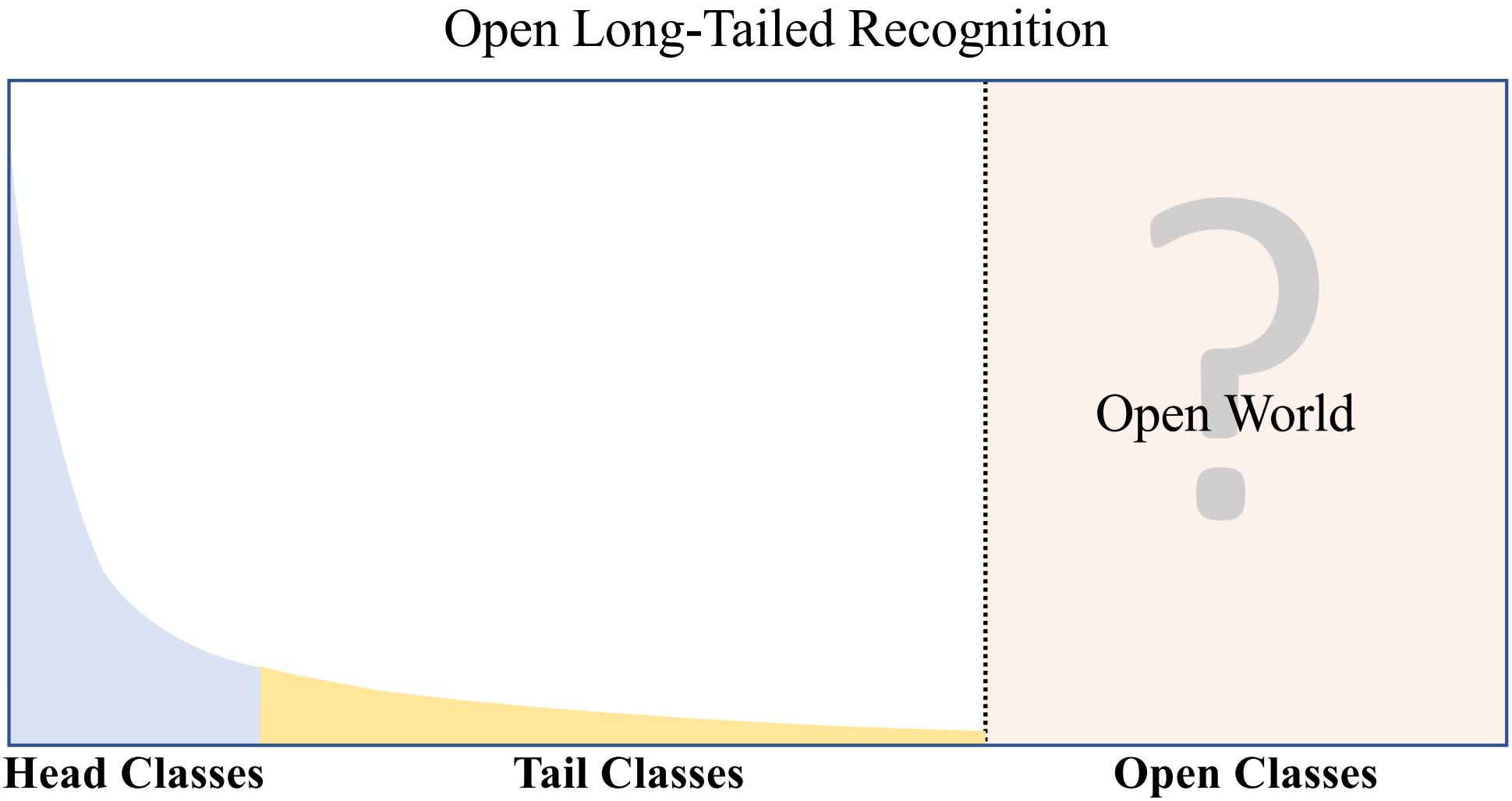
Places [Wang et al. 2017]

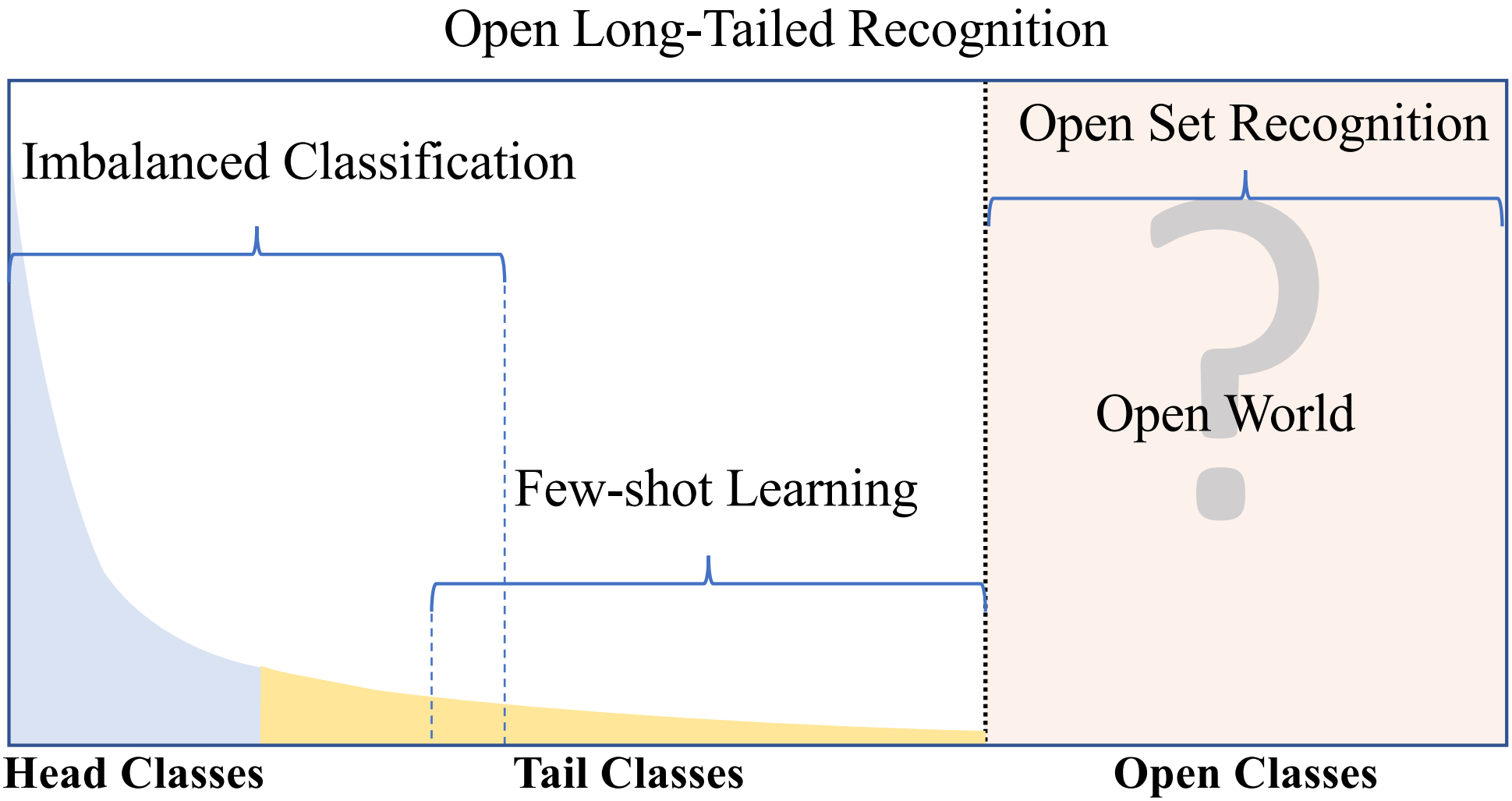


Species [Van Horn et al. 2019]



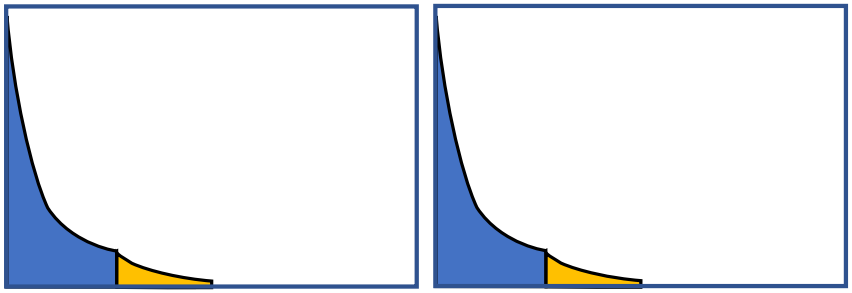
Actions [Zhang et al. 2019]


Open Long-Tailed Recognition
Open World
Head Classes
Tail Classes
Open Classes


Open Long-Tailed Recognition
Imbalanced Classification
Open Set Recognition
Open World
Few-shot Learning
Head Classes
Tail Classes
Open Classes

Imbalanced Classification

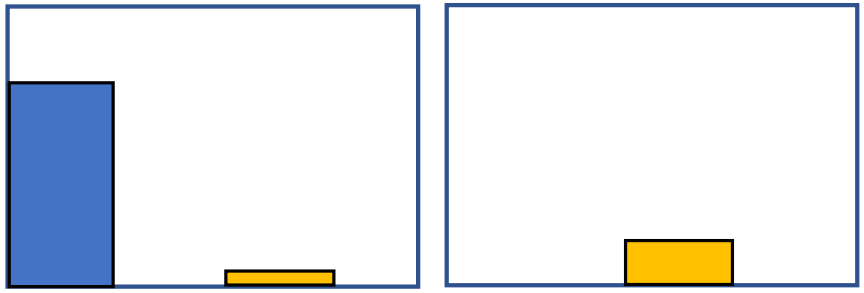
(metric learning, re-sampling, re-weighting)



Sensitivity to Novelty ✗

Few-Shot Learning

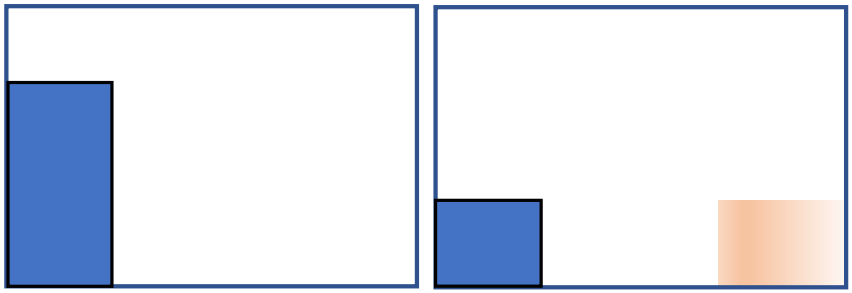
(meta learning, classifier dynamics)



Avoid Forgetting ✗

Open Set Recognition

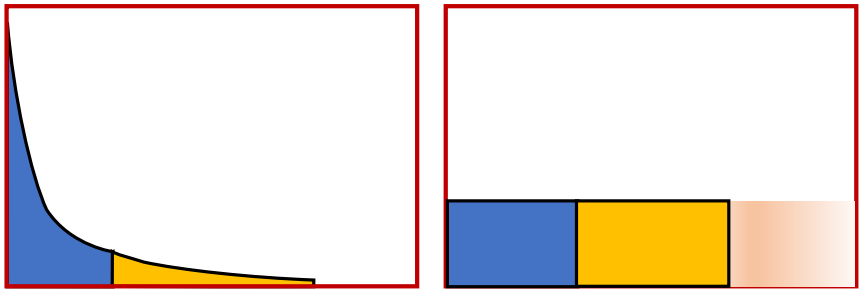
(distribution rectification, out-of-distribution detection)



Knowledge Transfer ✗

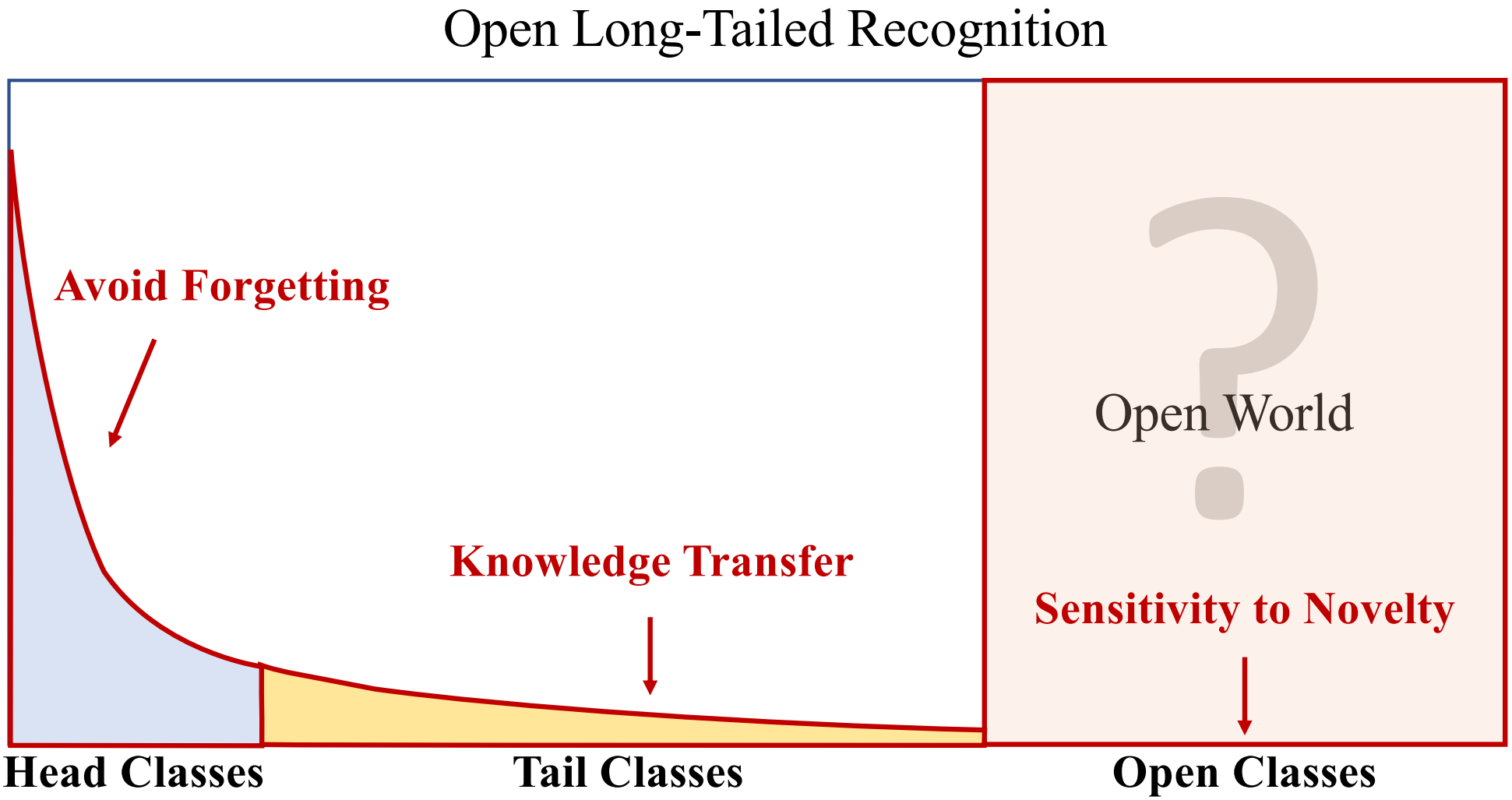
Open Long-Tailed Recognition

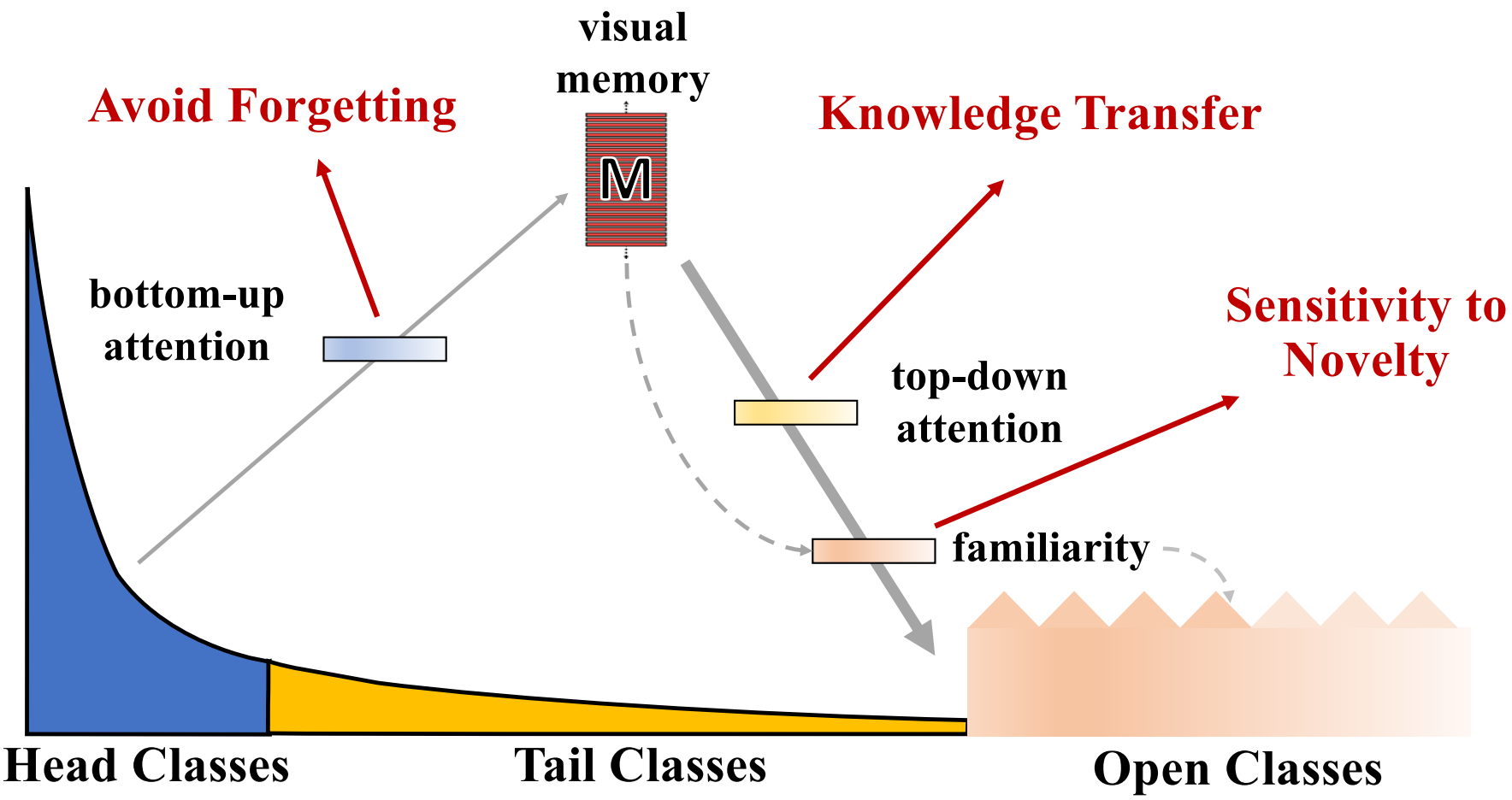
(dynamic meta-embedding)

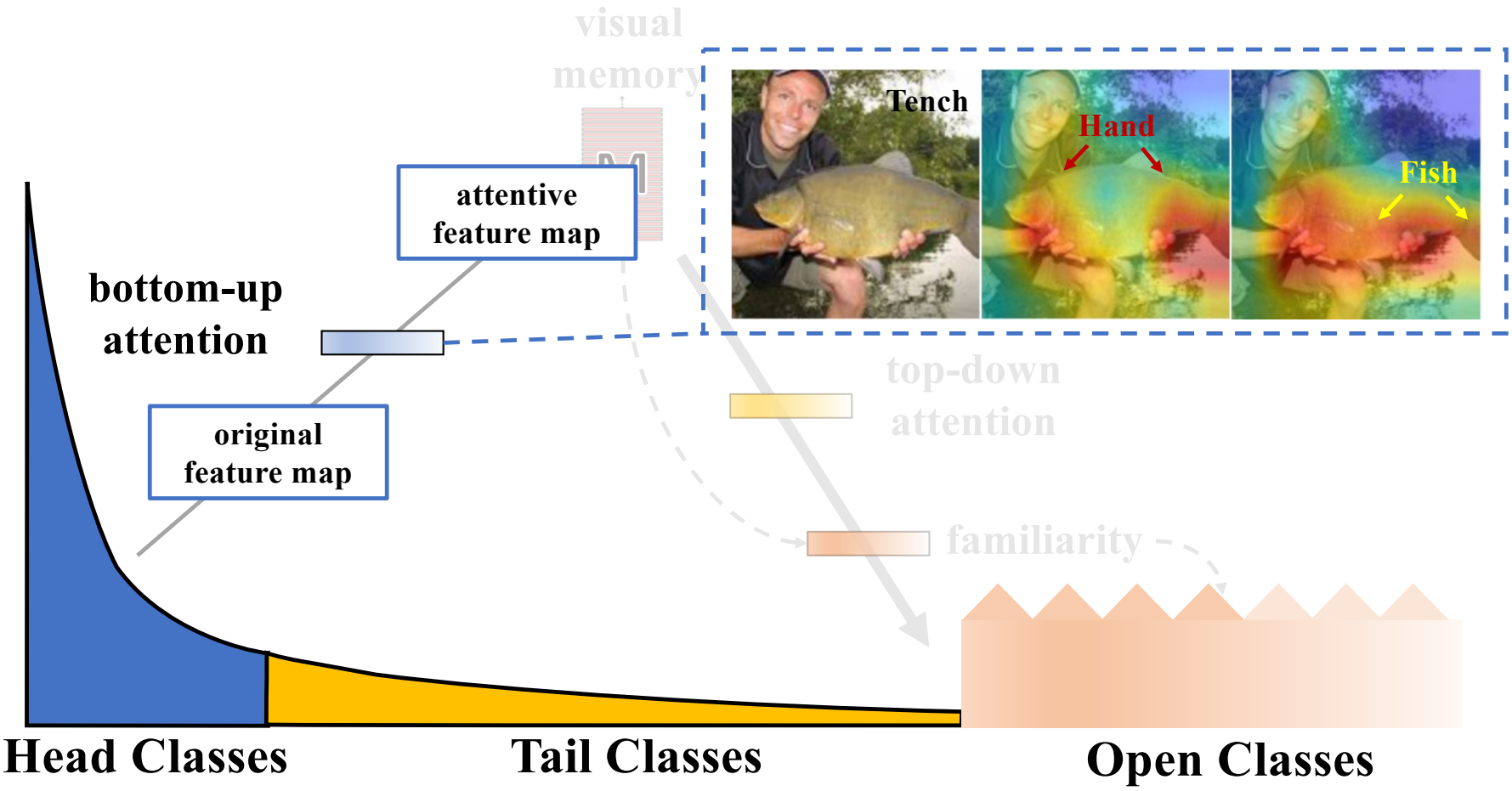


Knowledge Transfer ***Sensitivity to Novelty***

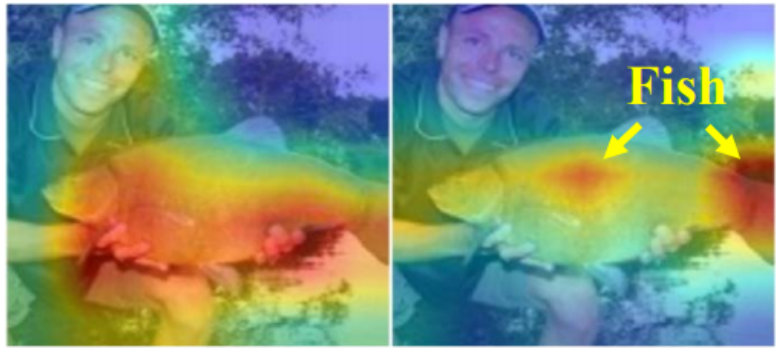
Avoid Forgetting


Open Long-Tailed Recognition
Avoid Forgetting
Knowledge Transfer
Open World
Sensitivity to Novelty
Head Classes
Tail Classes
Open Classes

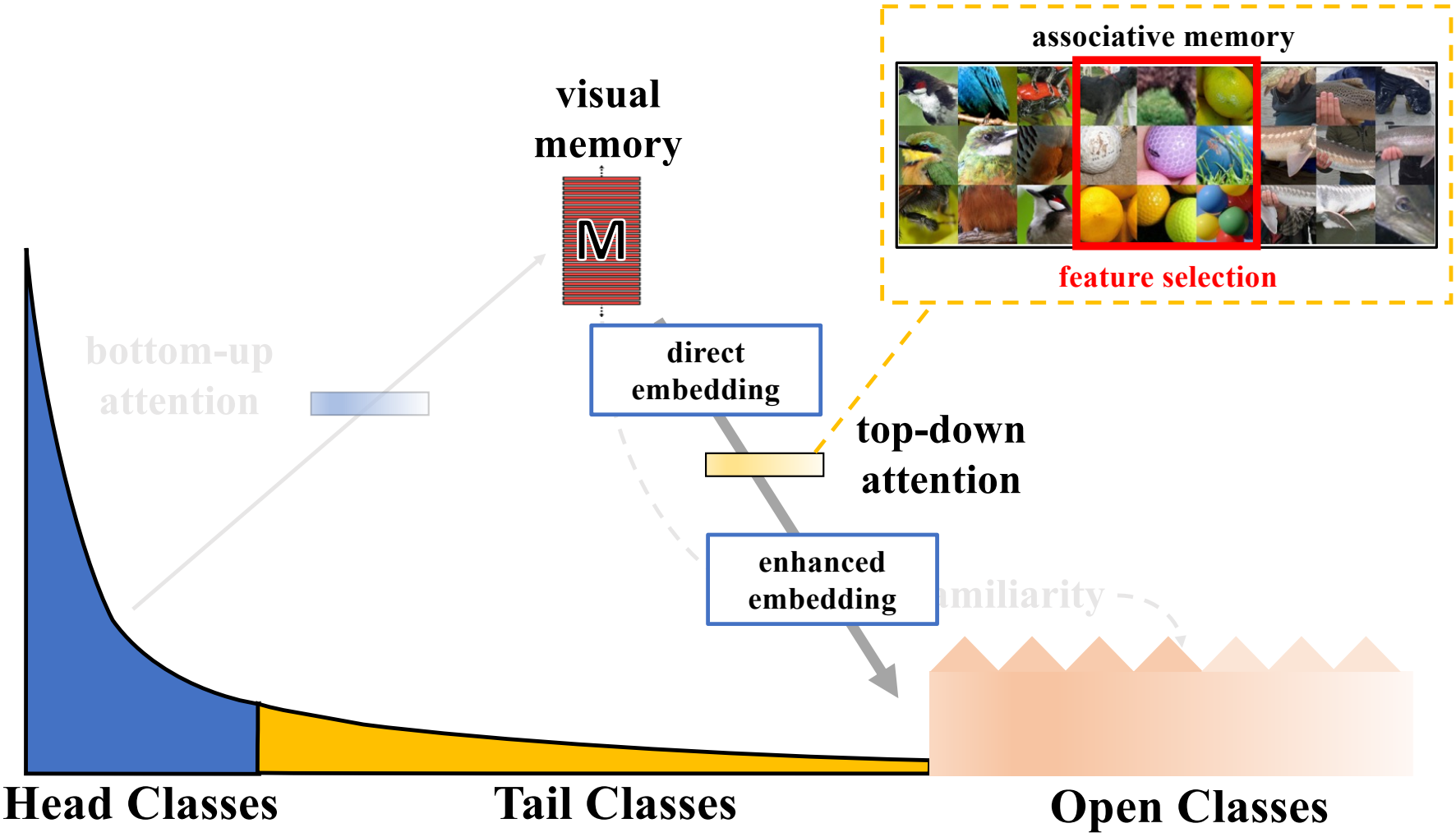

visual
memory
M
Avoid Forgetting
Knowledge Transfer
bottom-up
attention
top-down
attention
Sensitivity to
Novelty
familiarity
Head Classes
Tail Classes
Open Classes


visual
memory
Tench
Hand
Fish
attentive
feature map
bottom-up
attention
original
feature map
top-down
attention
familiarity
Head Classes
Tail Classes
Open Classes

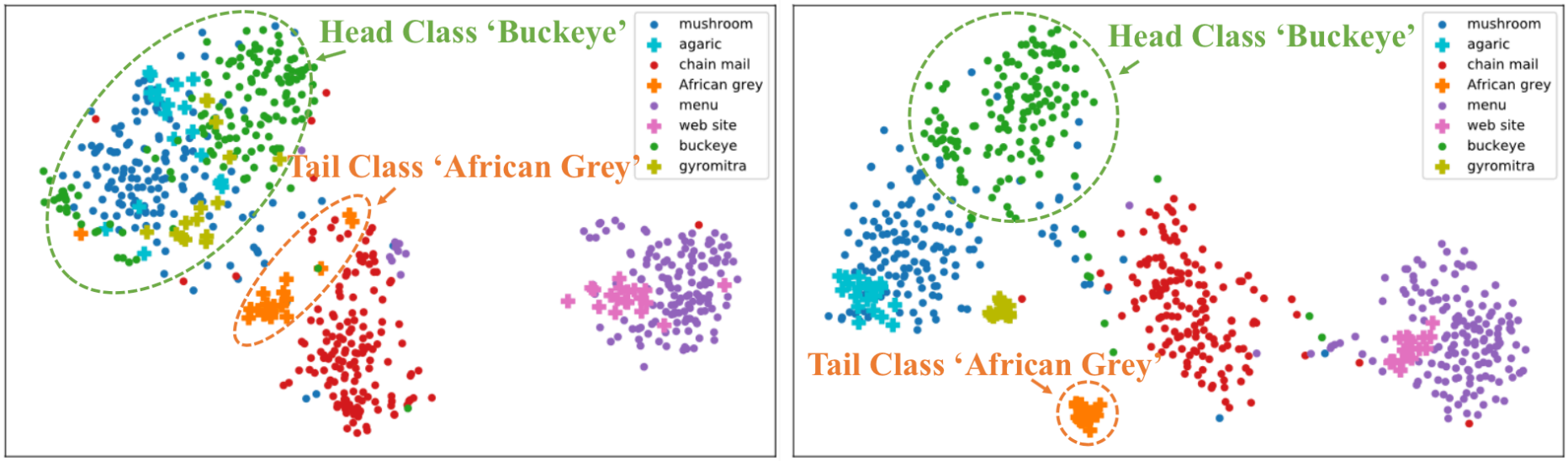
(b.1) Input Image (b.2) Feature Map of Plain ResNet Model



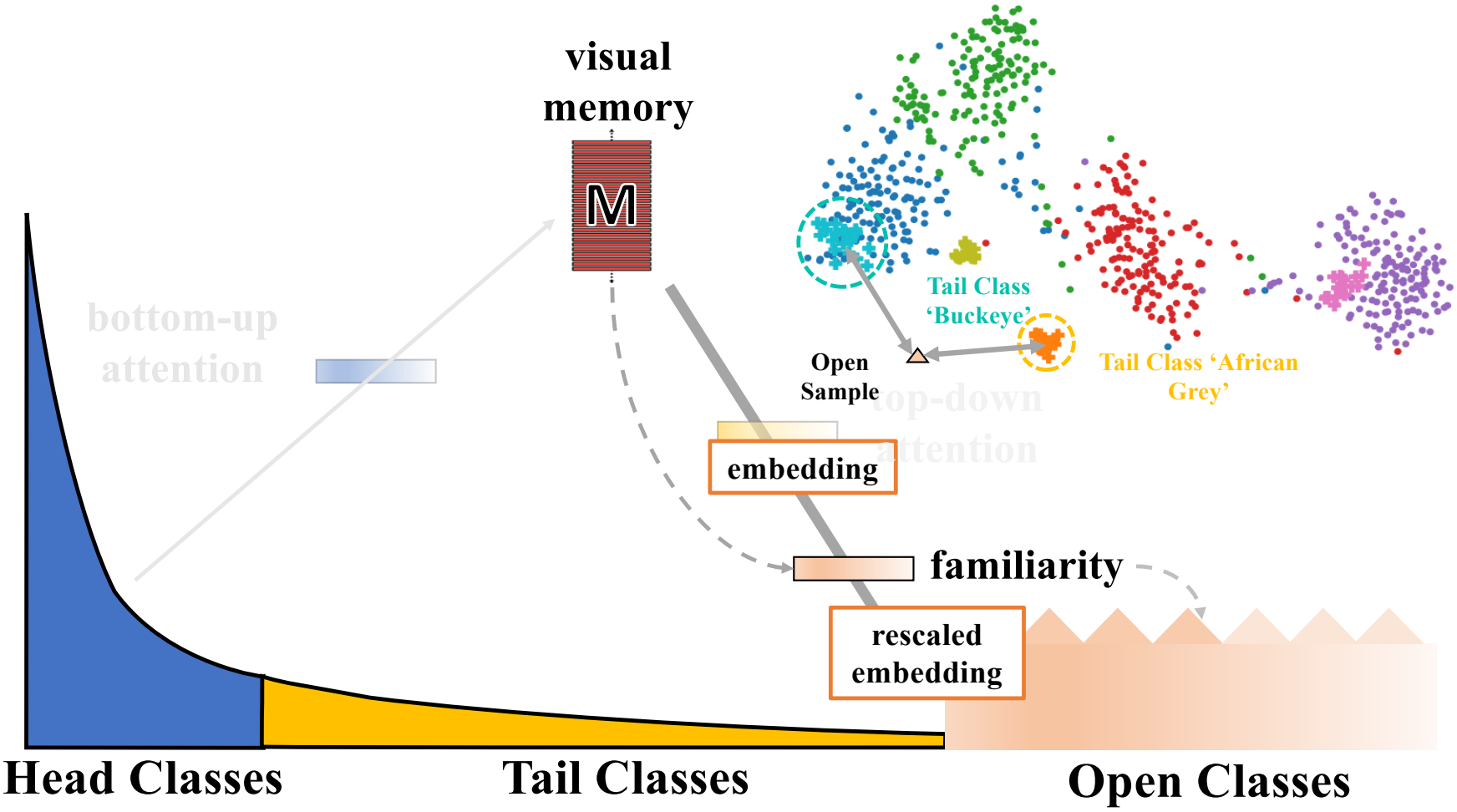
(b.3) Feature Map of Our Model (b.4) Modulated Attention

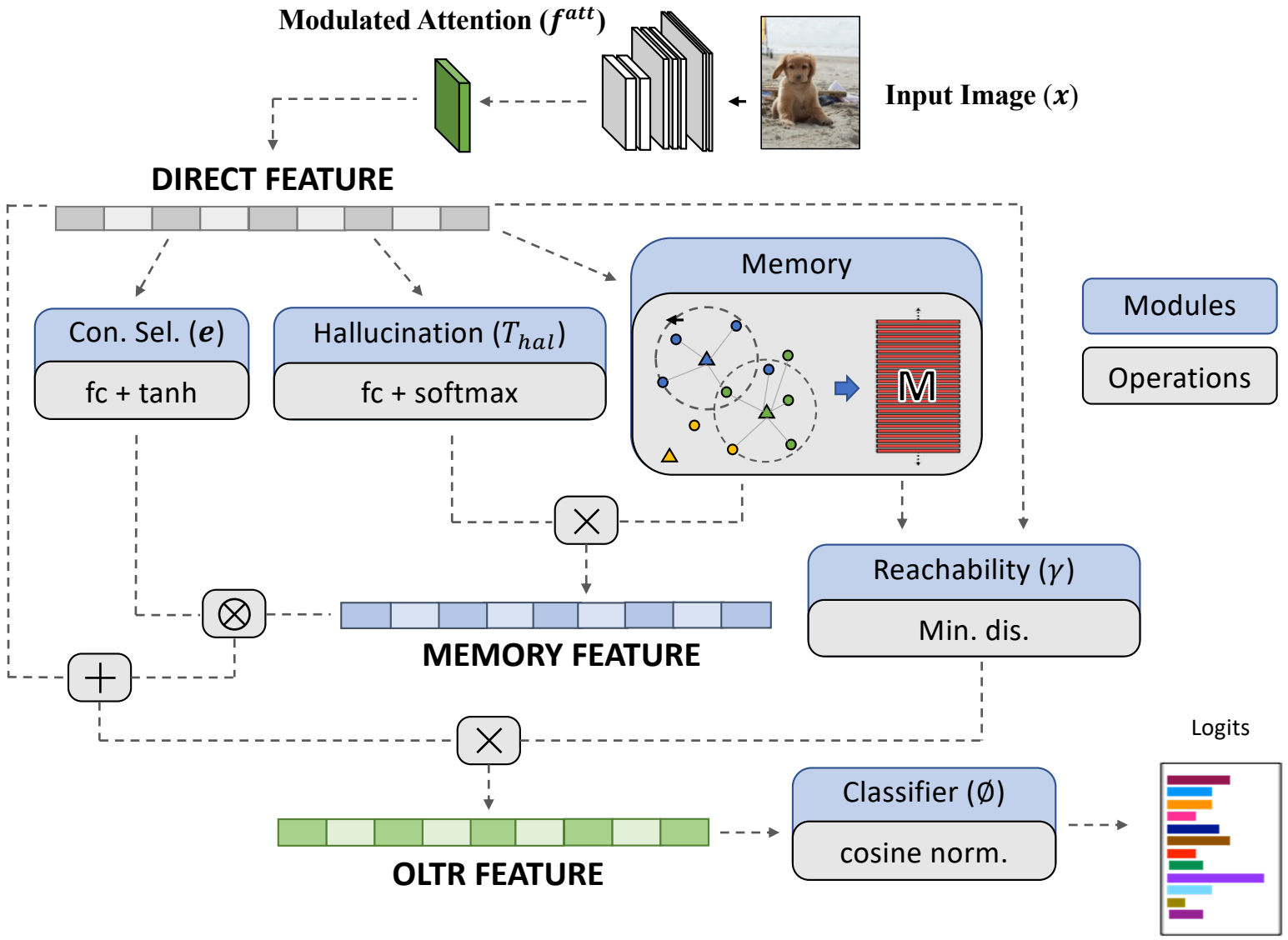

associative memory
feature selection
visual
memory
M
direct
embedding
top-down
attention
enhanced
embedding
bottom-up
attention
amiliarity
Head Classes
Tail Classes
Open Classes


Fly
Leef Beetle



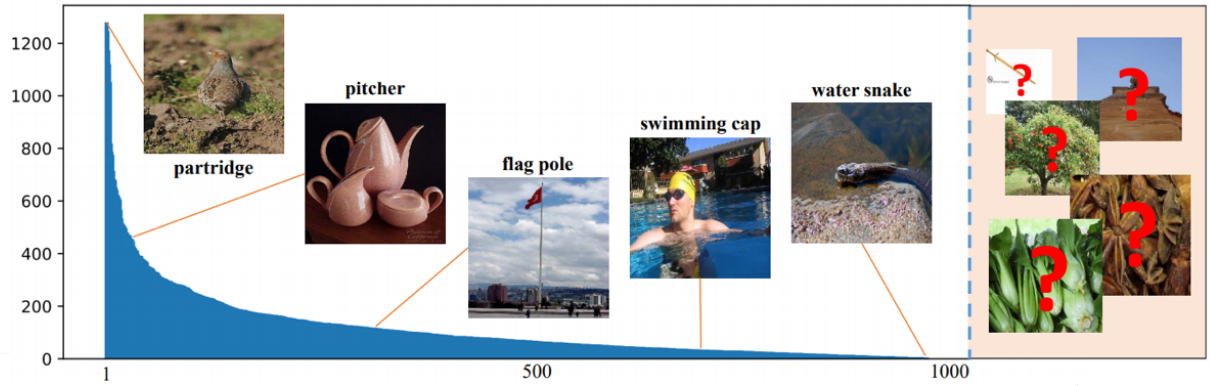
(a) Embedding of Plain ResNet Model (b) Embedding of Dynamic Meta-Embedding


visual
memory
M
bottom-up
attention
Tail Class
'Buckeye'
Open
Sample
Tail Class 'African
Grey'
top-down
attention
embedding
familiarity
rescaled
embedding
Head Classes
Tail Classes
Open Classes


Modulated Attention (f^{att})
Input Image (x)
DIRECT FEATURE
Memory
M
Modules
Operations
Con. Sel. (e)
fc + tanh
Hallucination (T_{hal})
fc + softmax
Reachability (γ)
Min. dis.
MEMORY FEATURE
OLTR FEATURE
Classifier ($\emptyset$)
cosine norm.
Logits

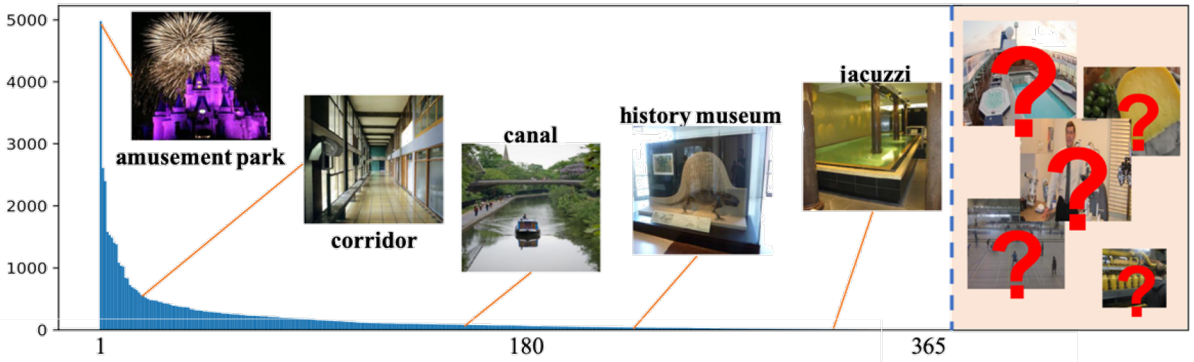
ImageNet-LT Benchmark

Absolute Performance Gain: ~20%



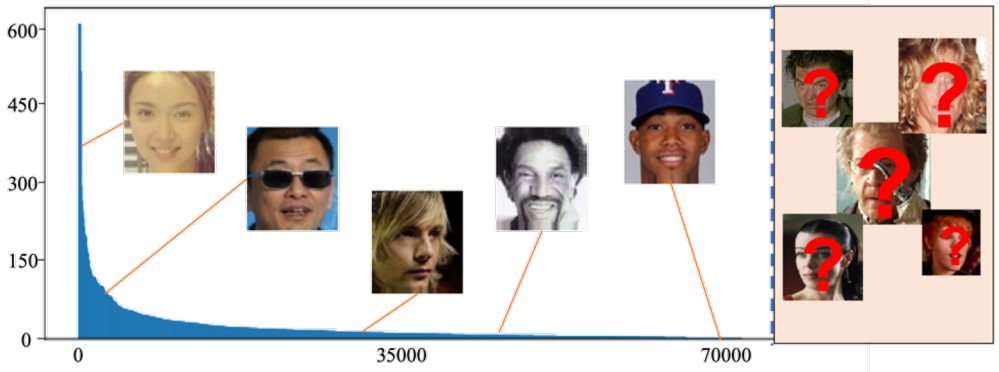
Places-LT Benchmark

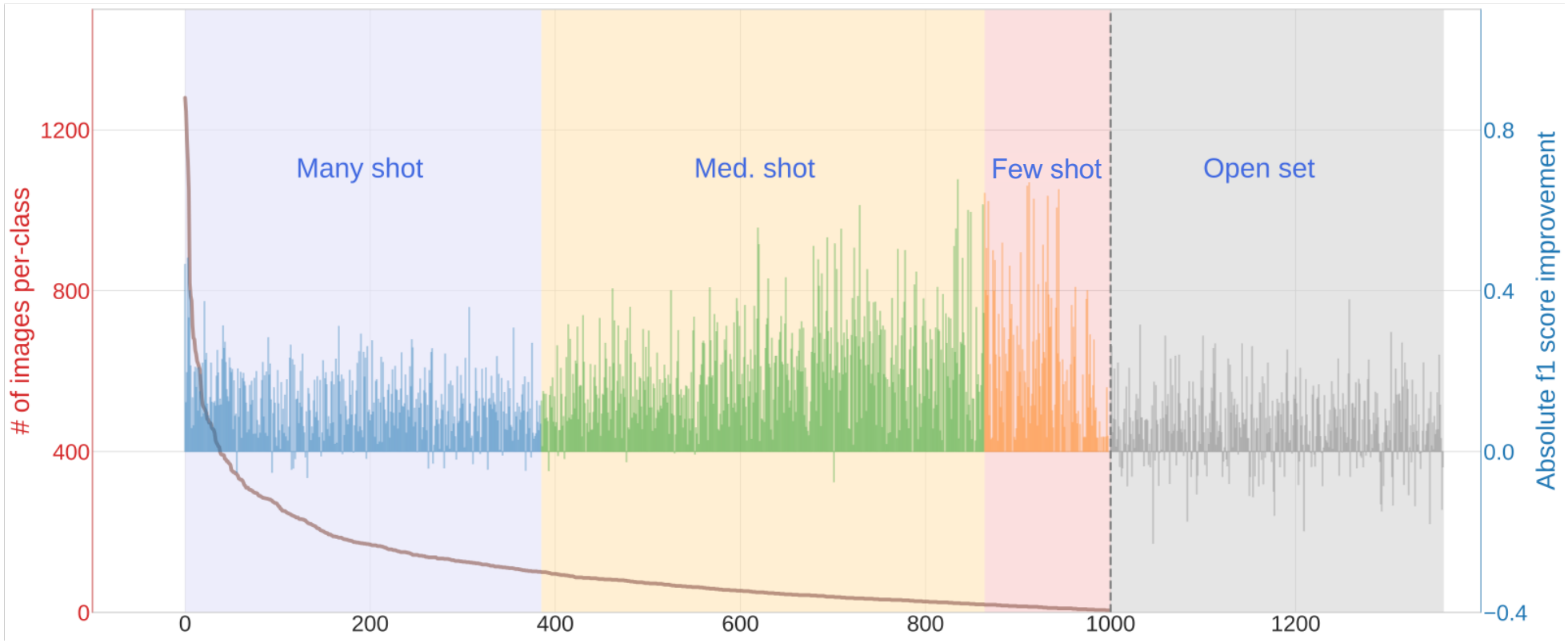
Absolute Performance Gain: ~10%



MS1M-LT Benchmark

Absolute Performance Gain: ~2%




# of images per-class
1200
800
400
0
Many shot
Med. shot
Few shot
Open set
0
200
400
600
800
1000
1200
0.8
0.4
0.0
−0.4
Absolute f1 score improvement

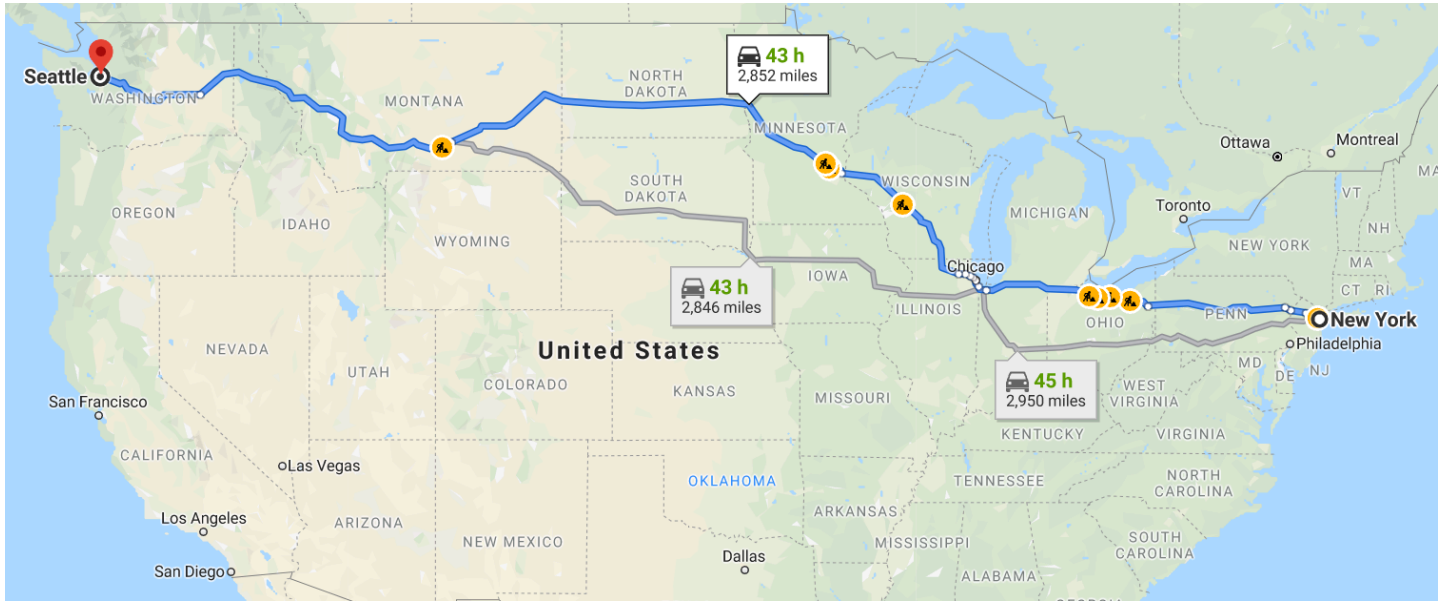
Open compound domain adaptation

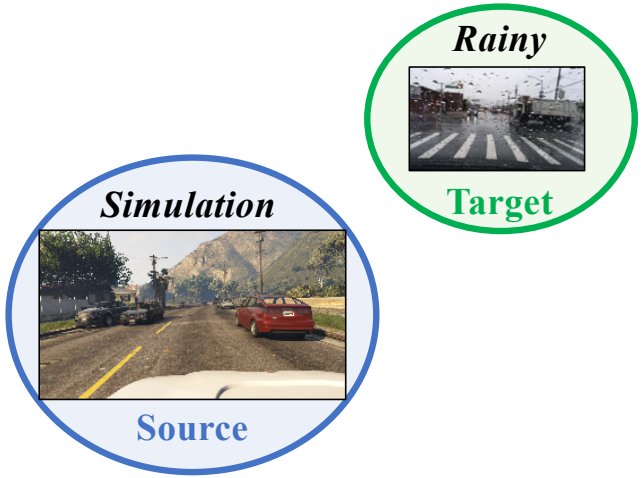
Ziwei Liu*, Zhongqi Miao*, Xingang Pan, Xiaohang Zhan, Dahua Lin,
Stella X. Yu, Boqing Gong
The Chinese University of Hong Kong & UC Berkeley / ICSI &
Google Inc.

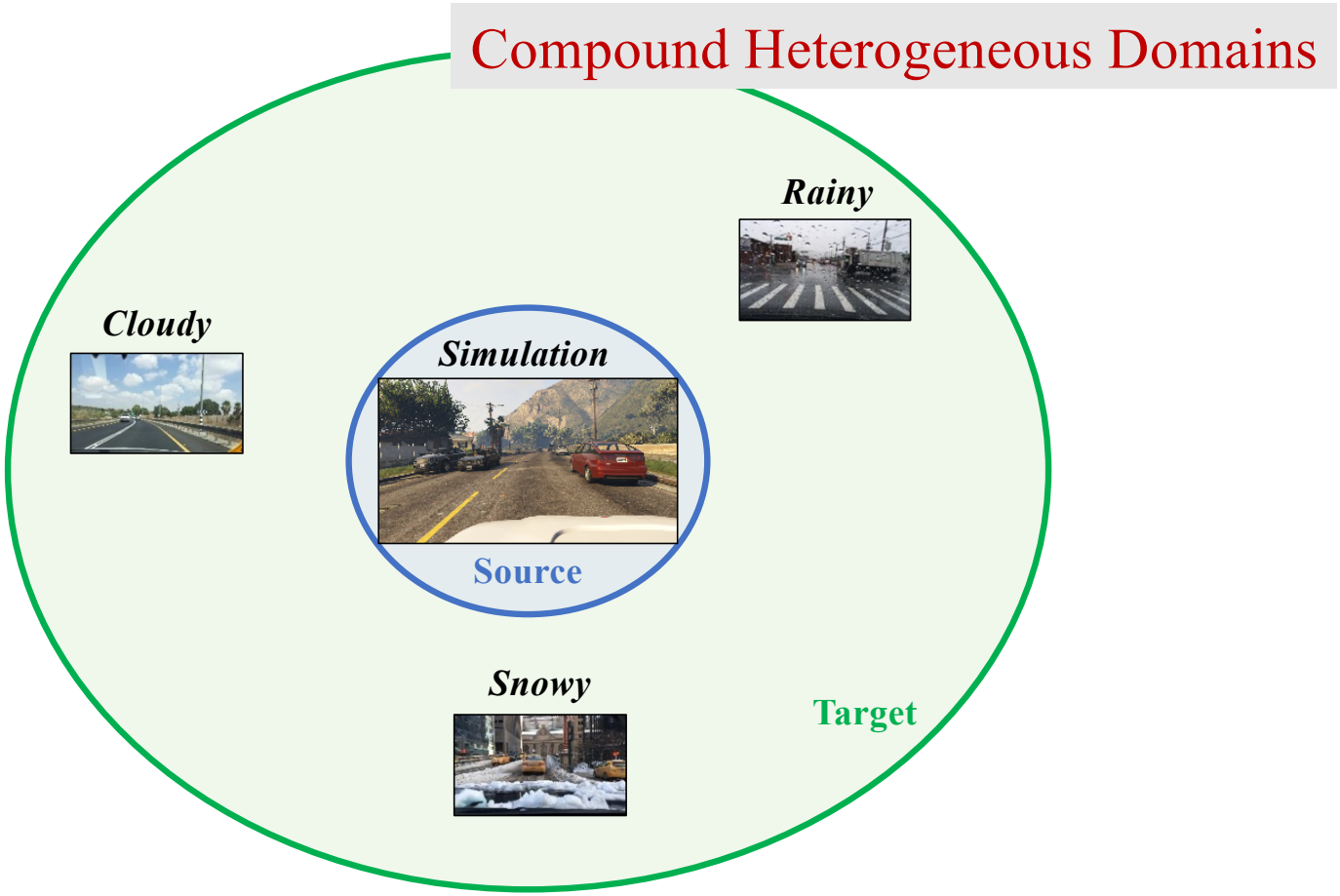
CVPR, 2020, oral

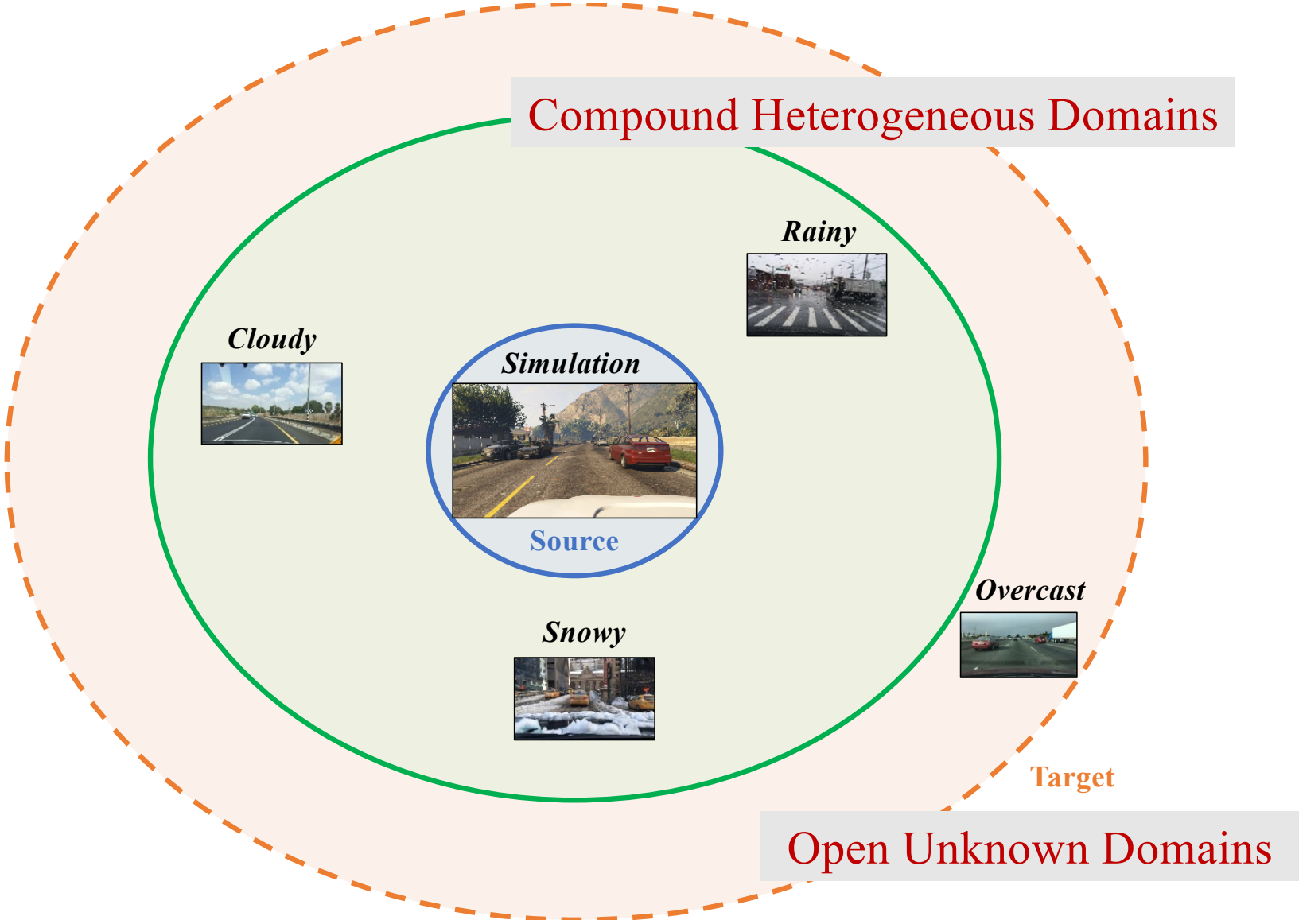
Simulation

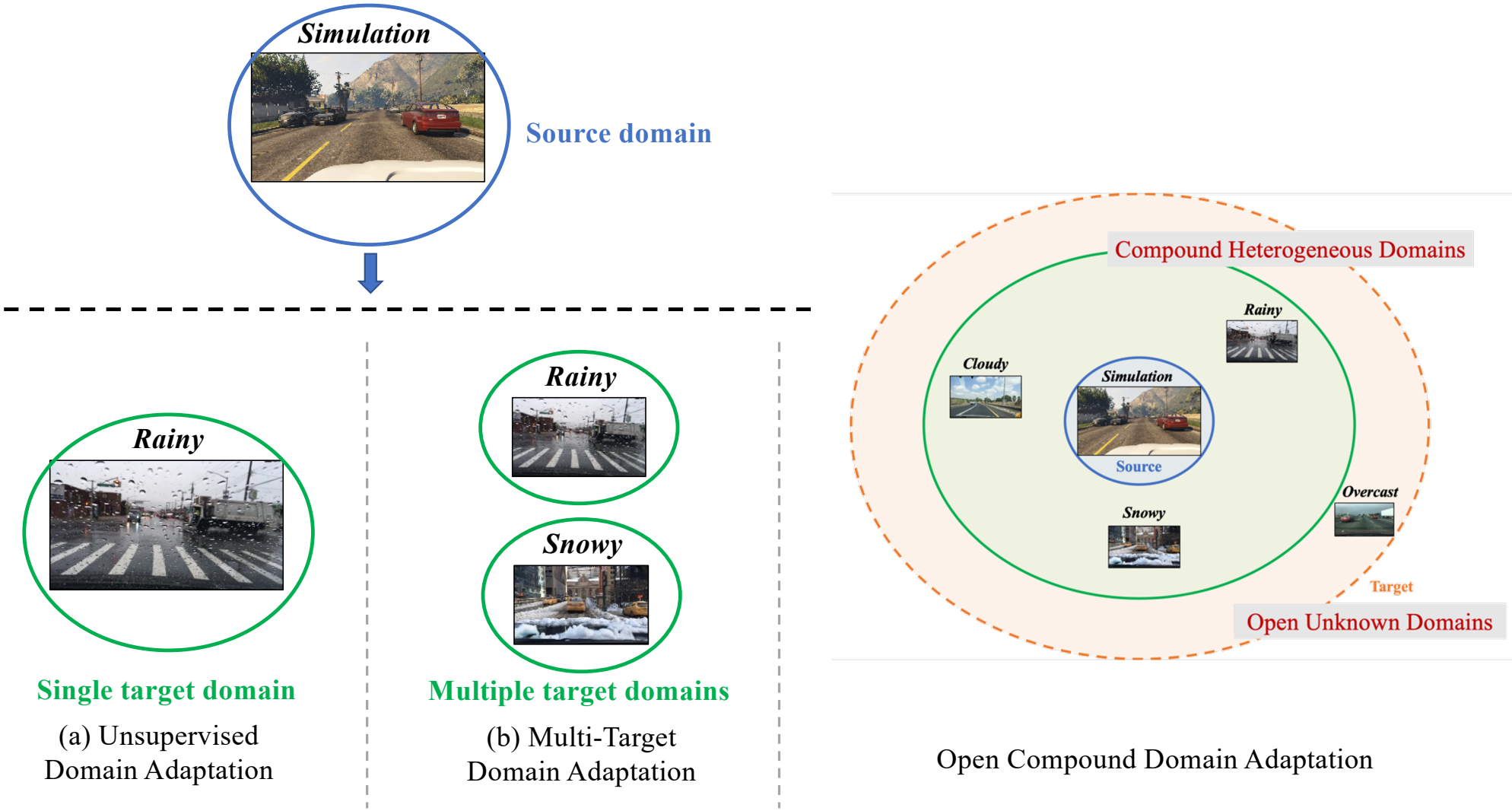
Open World Driving Conditions

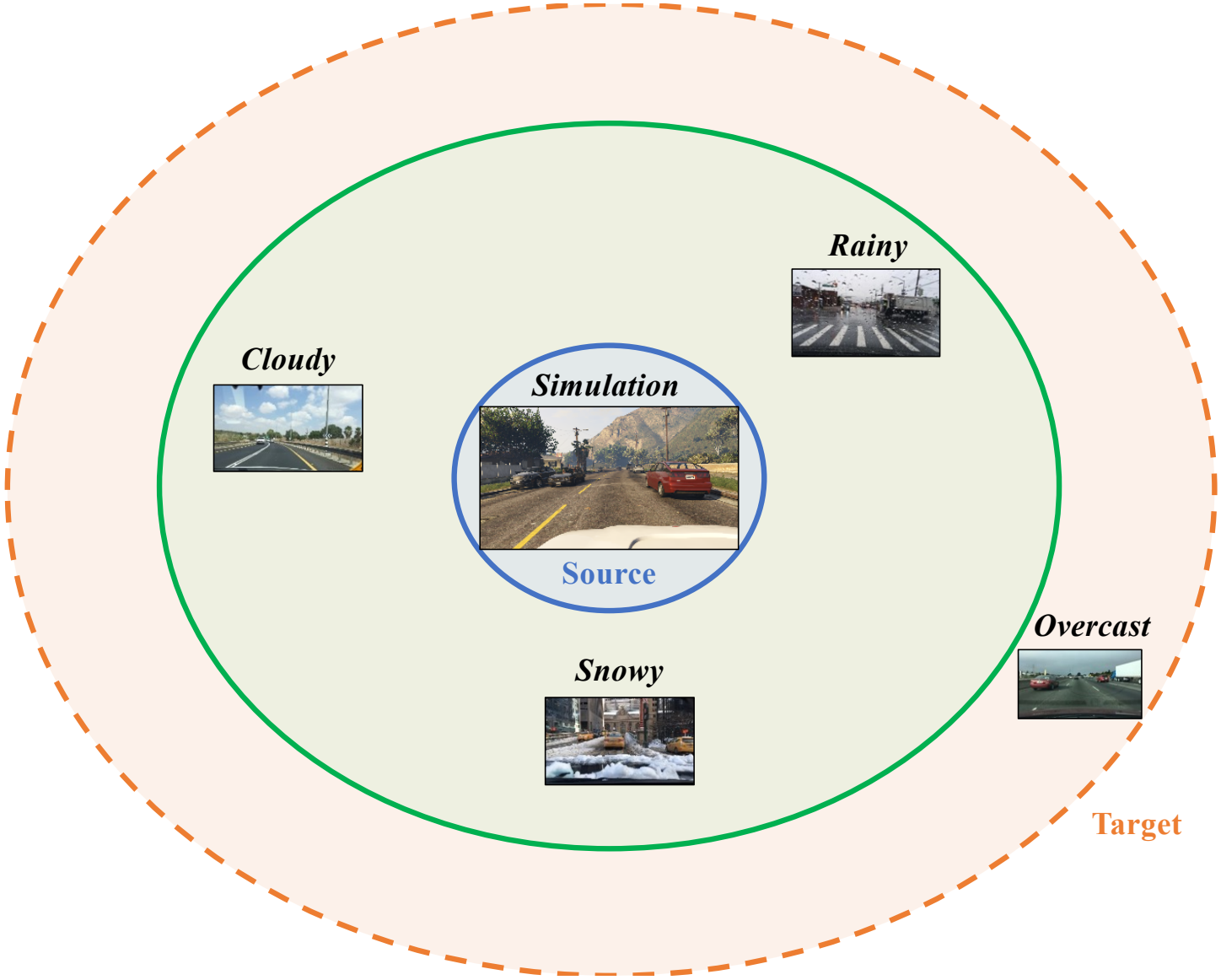


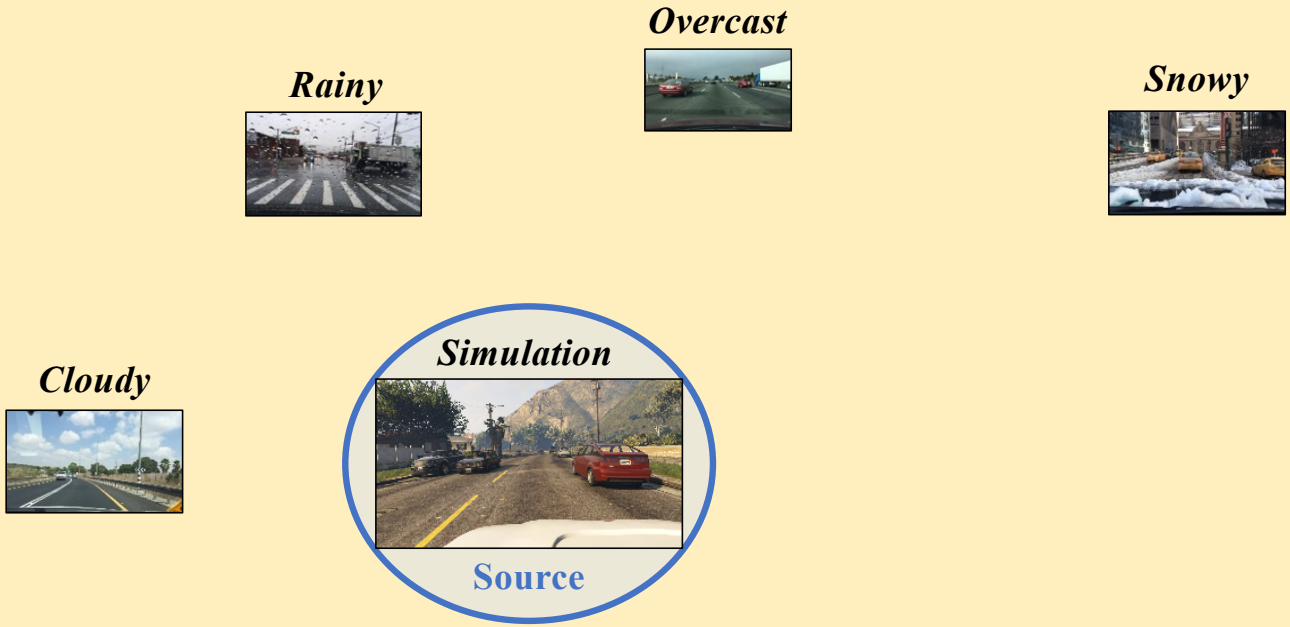

Rainy
Target
Simulation
Source


Compound Heterogeneous Domains
Rainy
Cloudy
Simulation
Source
Snowy
Target


Compound Heterogeneous Domains
Rainy
Cloudy
Simulation
Source
Overcast
Snowy
Target
Open Unknown Domains


Simulation
Source domain
Rainy
Single target domain
(a) Unsupervised Domain Adaptation
Rainy
Snowy
Multiple target domains
(b) Multi-Target Domain Adaptation
Compound Heterogeneous Domains
Cloudy
Simulation
Source
Rainy
Overcast
Snowy
Target
Open Unknown Domains
Open Compound Domain Adaptation


Rainy
Cloudy
Simulation
Source
Overcast
Snowy
Target


Overcast
Rainy
Snowy
Cloudy
Simulation
Source

Foggy

After-rain

Thunderstorm

Snowy

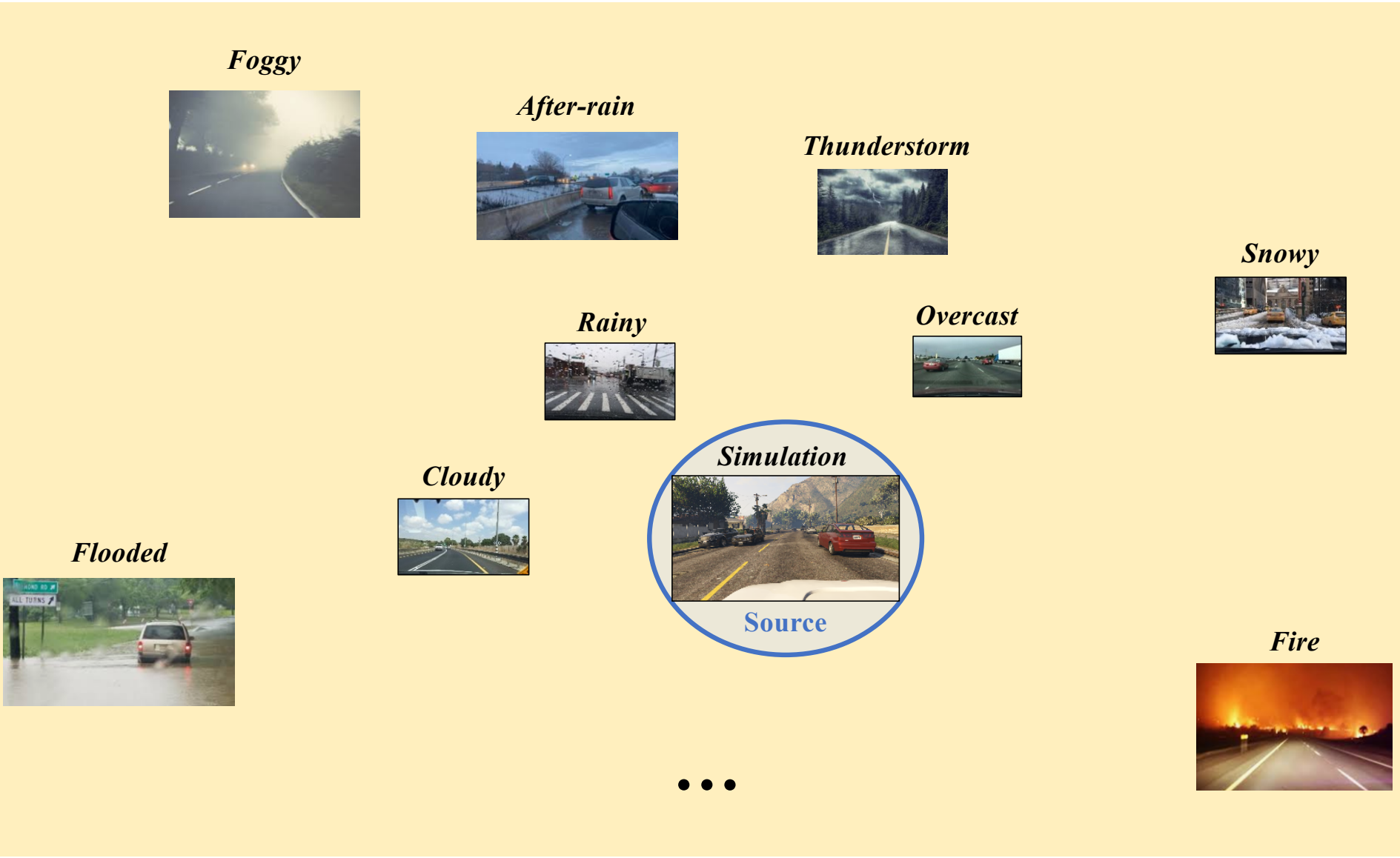
Rainy

Overcast

Cloudy

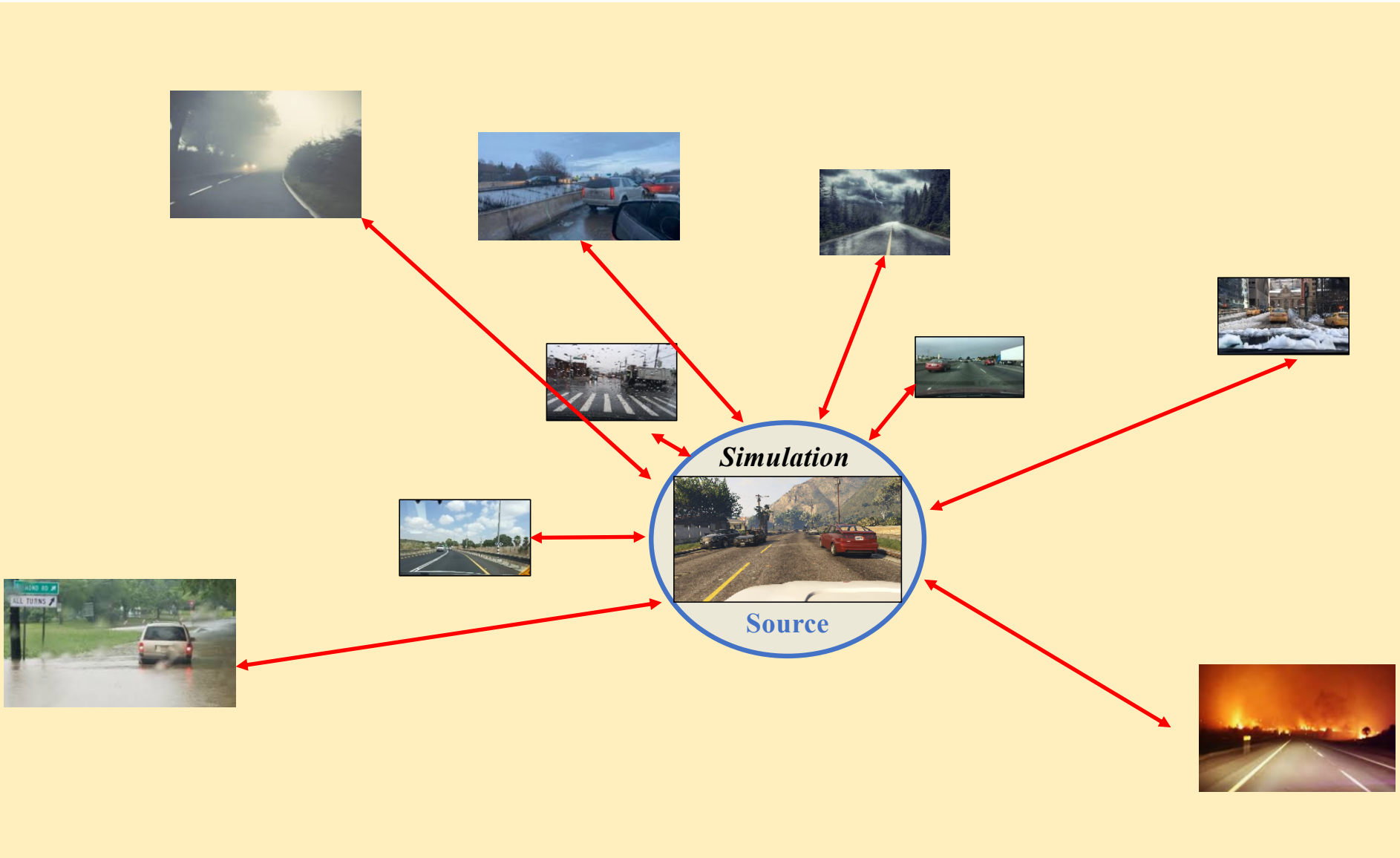
Simulation

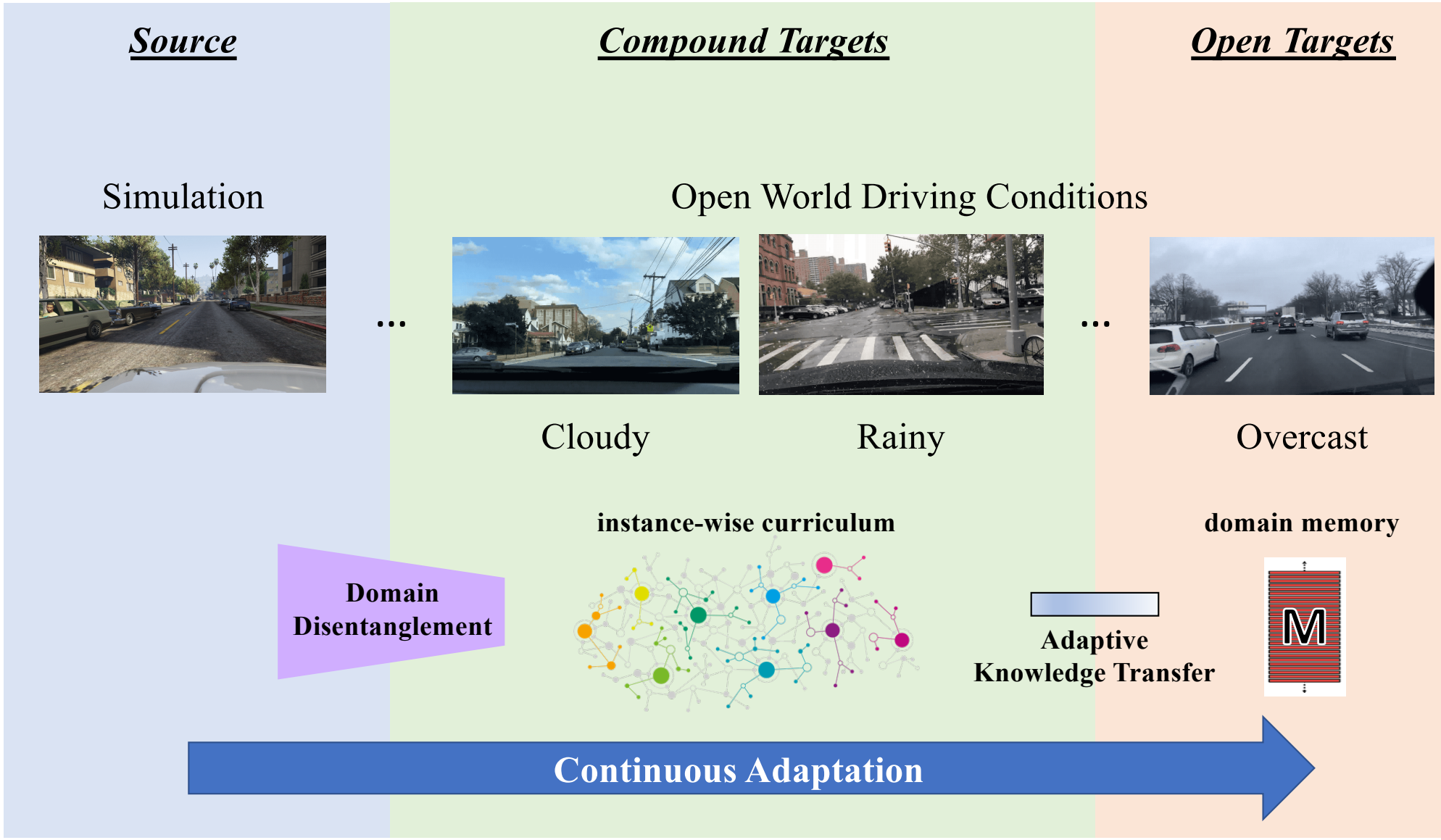
Source

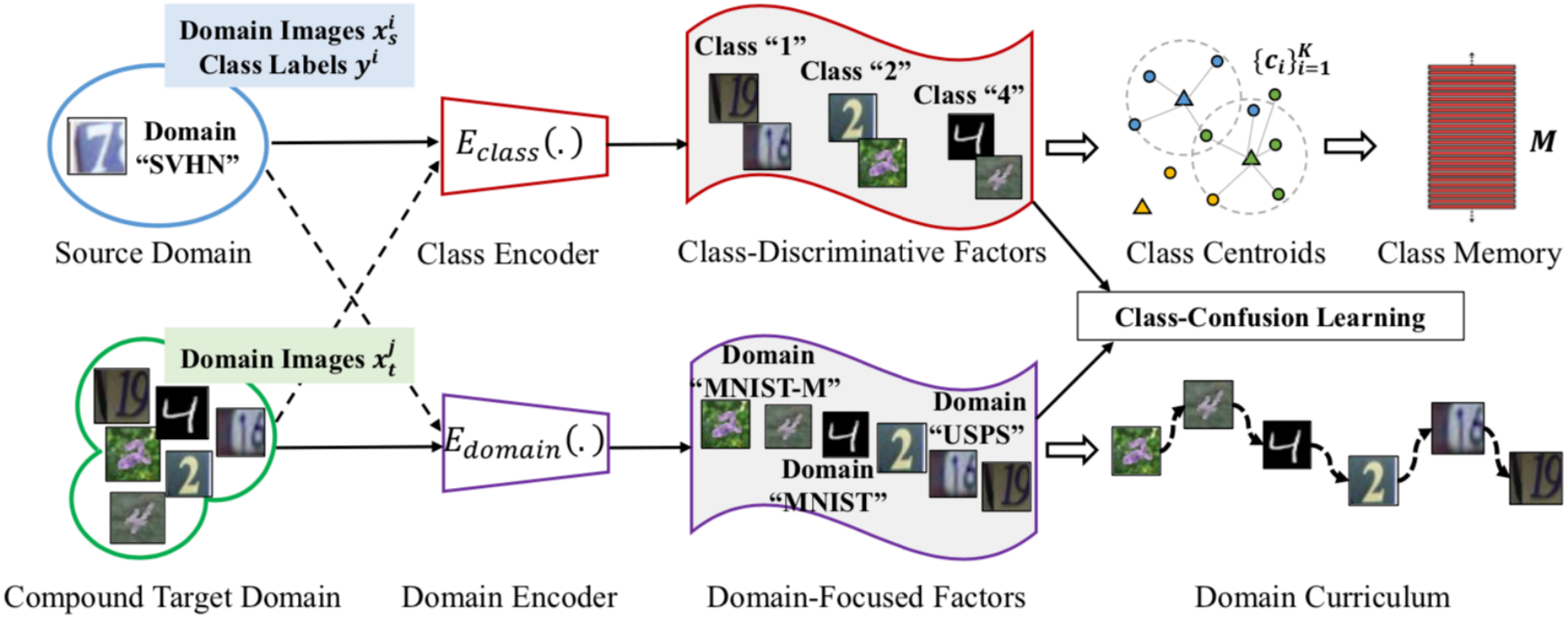

Foggy
After-rain
Thunderstorm
Snowy
Rainy
Overcast
Simulation
Cloudy
Source
Flooded
Fire

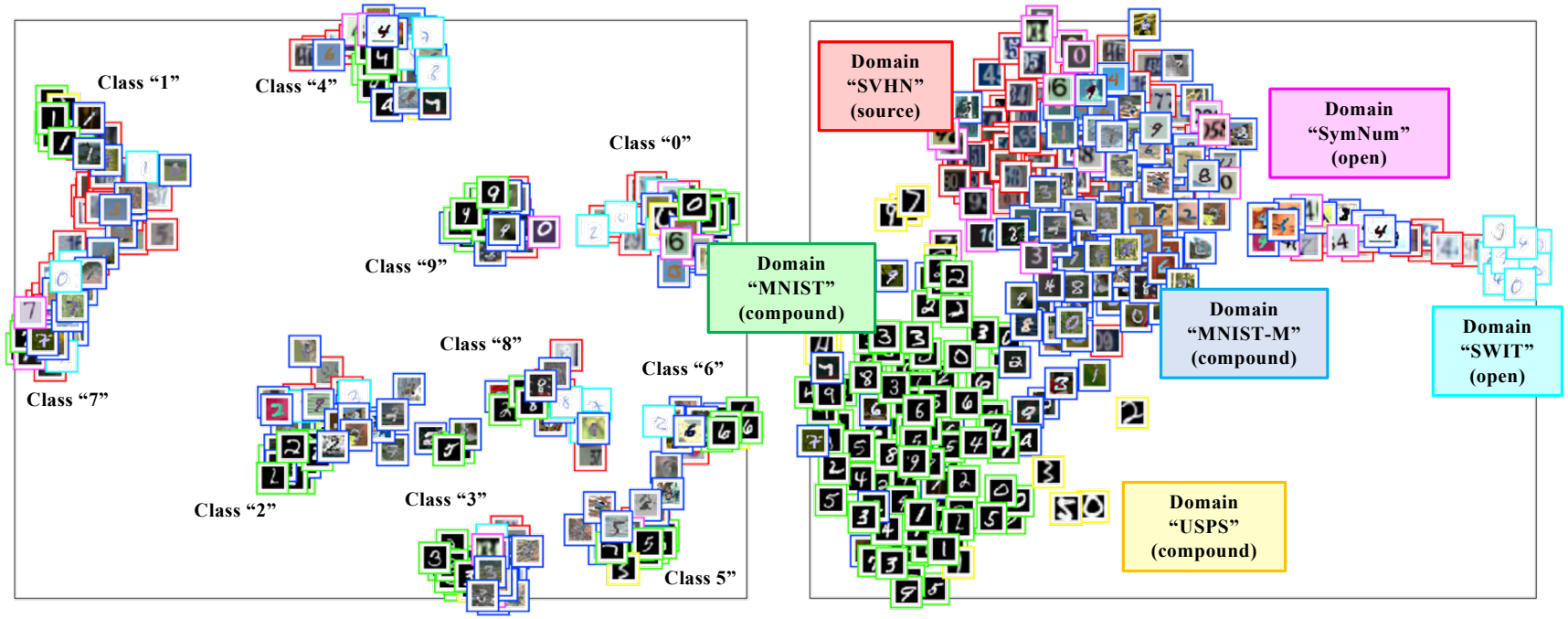
Simulation

Source


Simulation
Source

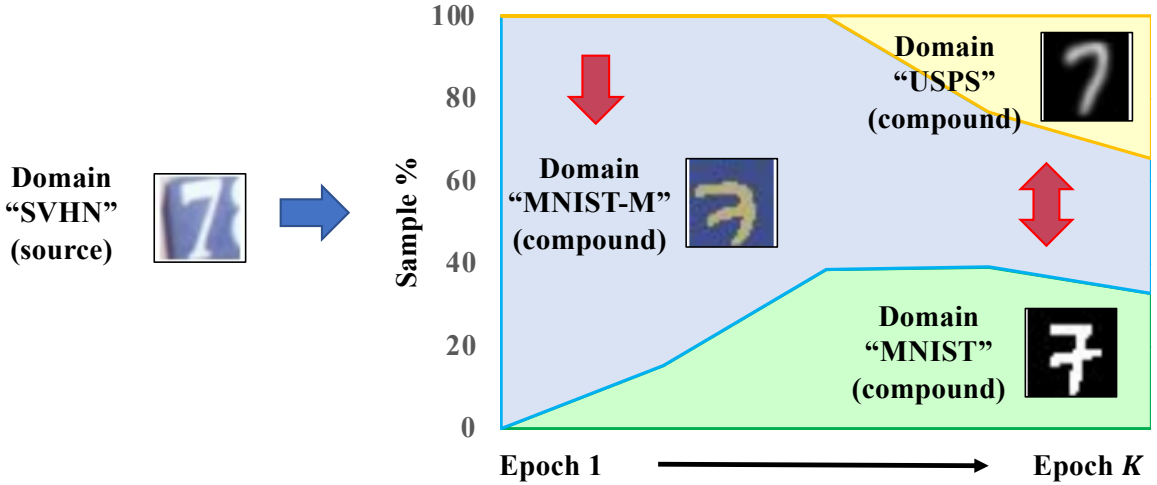

Source
Compound Targets
Open Targets
Simulation
Open World Driving Conditions
...
...
Cloudy
Rainy
Overcast
instance-wise curriculum
domain memory
Domain Disentanglement
Adaptive Knowledge Transfer
M
Continuous Adaptation

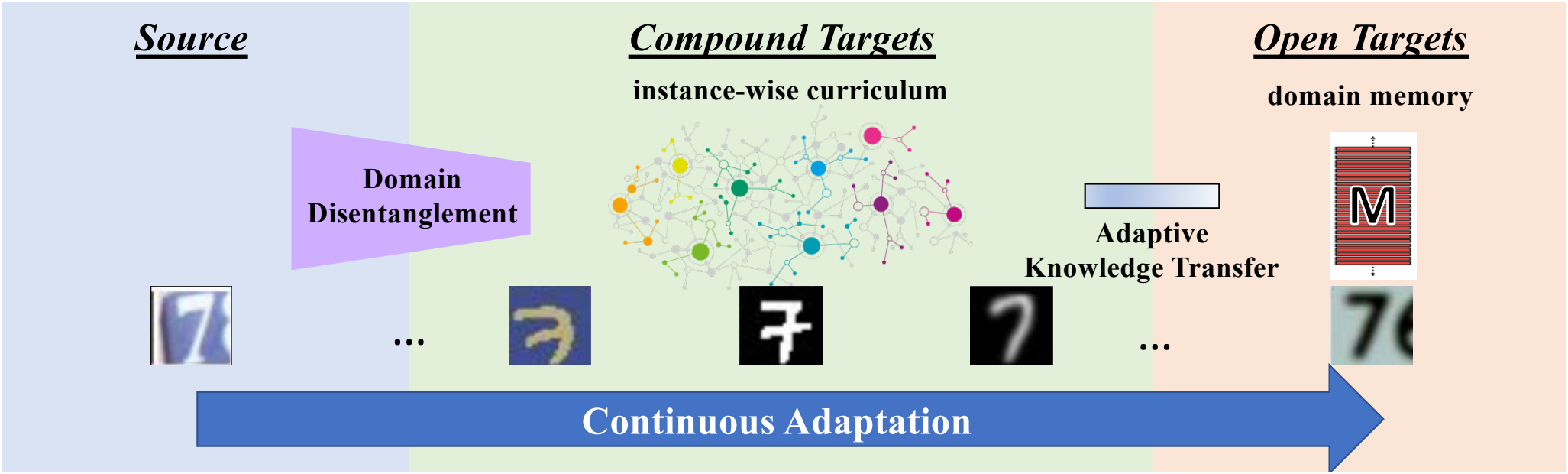

Domain Images x_s^i
Class Labels y^i
Domain "SVHN"
Source Domain
$E_{class}(.)$
Class Encoder
Class "1"
Class "2"
Class "4"
Class-Discriminative Factors
$\{c_i\}_{i=1}^K$
Class Centroids
M
Class Memory
Class-Confusion Learning
Domain Images x_t^j
Compound Target Domain
$E_{domain}(.)$
Domain Encoder
Domain "MNIST-M"
Domain "USPS"
Domain "MNIST"
Domain-Focused Factors
Domain Curriculum

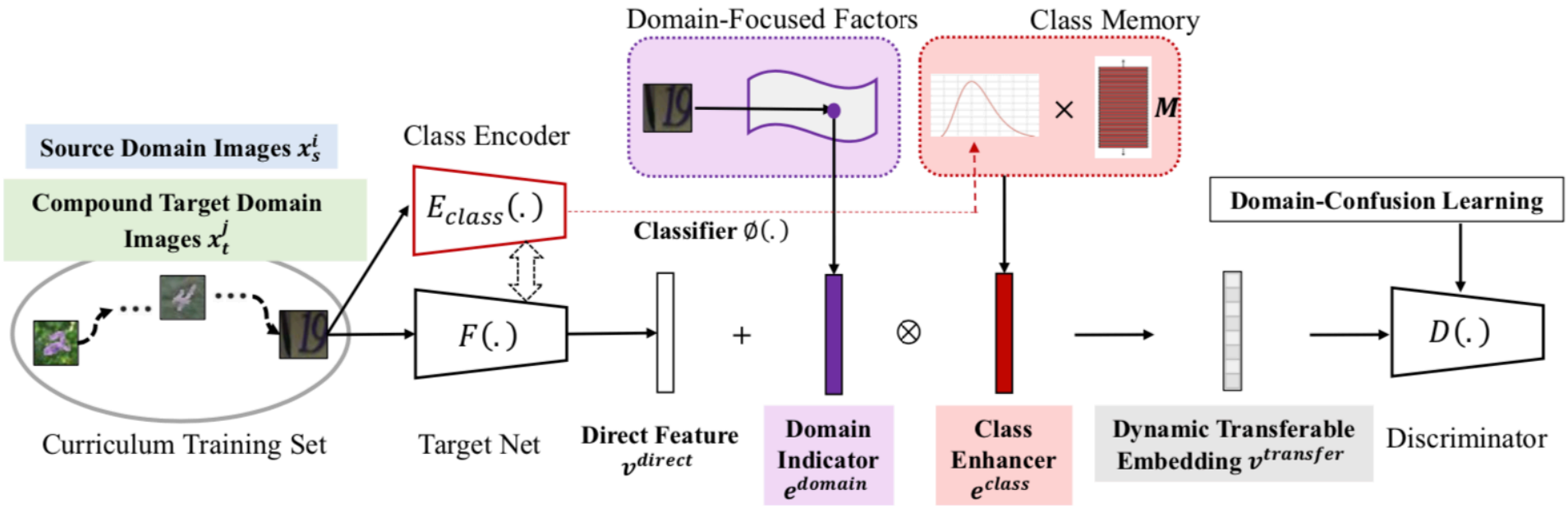

Class “1”
Class “4”
Class “0”
Class “9”
Class “8”
Class “6”
Class “7”
Class “2”
Class “3”
Class “5”
Domain “MNIST” (compound)
Domain “SVHN” (source)
Domain “SymNum” (open)
Domain “MNIST-M” (compound)
Domain “SWIT” (open)
Domain “USPS” (compound)


svhn_bal
mnist
mnistm
usps

Curriculum according to Domain Characteristics

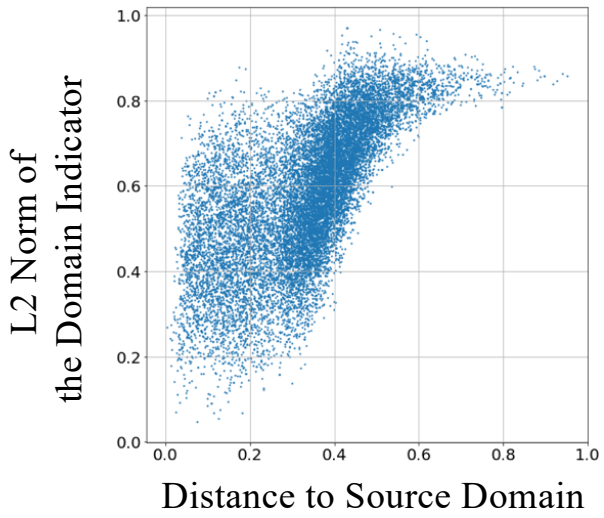


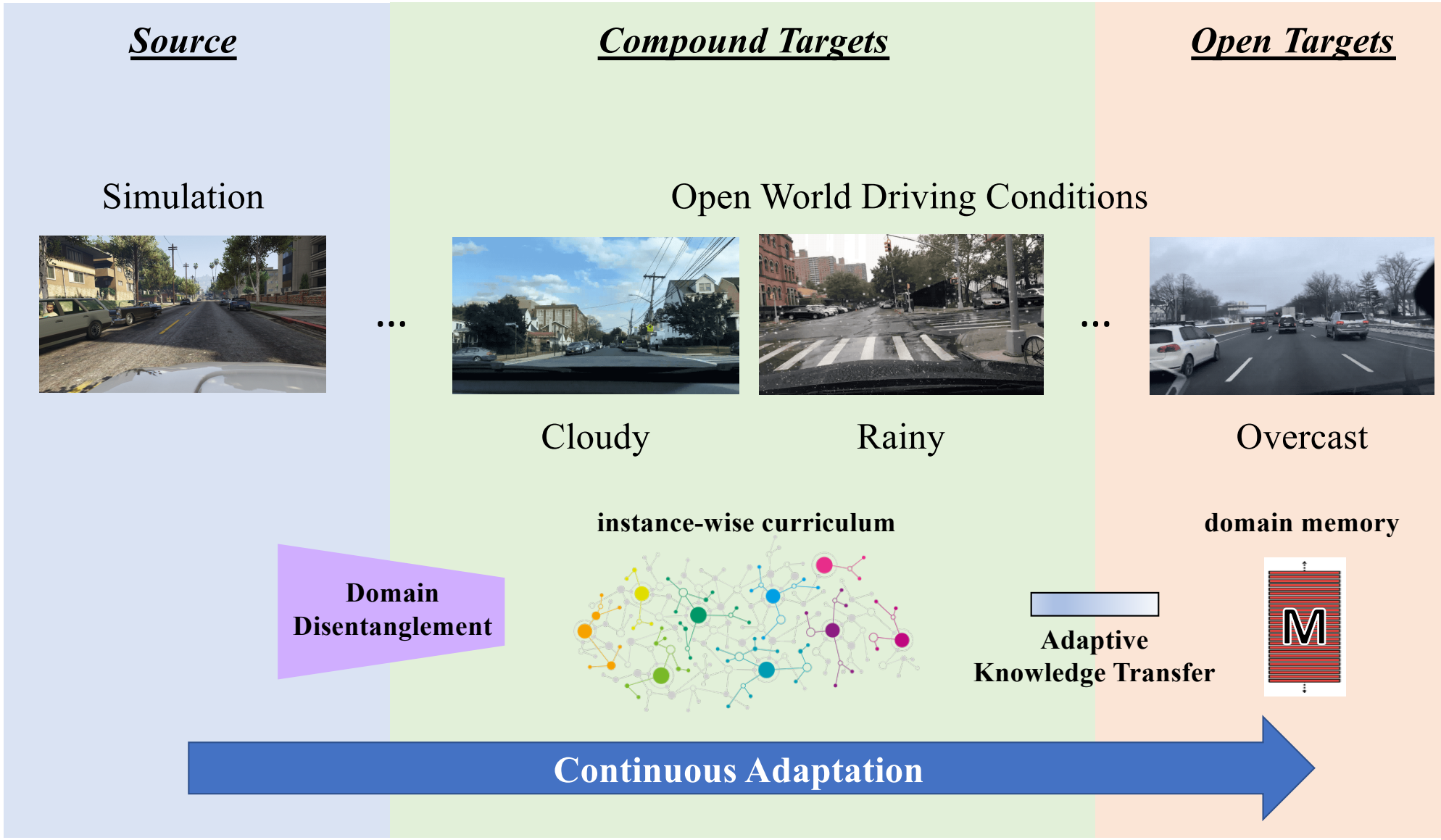



Domain-Focused Factors
Class Memory
Class Encoder
Source Domain Images x_s^i
Compound Target Domain Images x_t^j
$E_{class}(.)$
Classifier $\emptyset(.)$
M
Domain-Confusion Learning
$F(.)$
$D(.)$
Curriculum Training Set
Target Net
Direct Feature v^{direct}
Domain Indicator e^{domain}
Class Enhancer e^{class}
Dynamic Transferable Embedding $v^{transfer}$
Discriminator

Memory-Augmented Domain Indicator

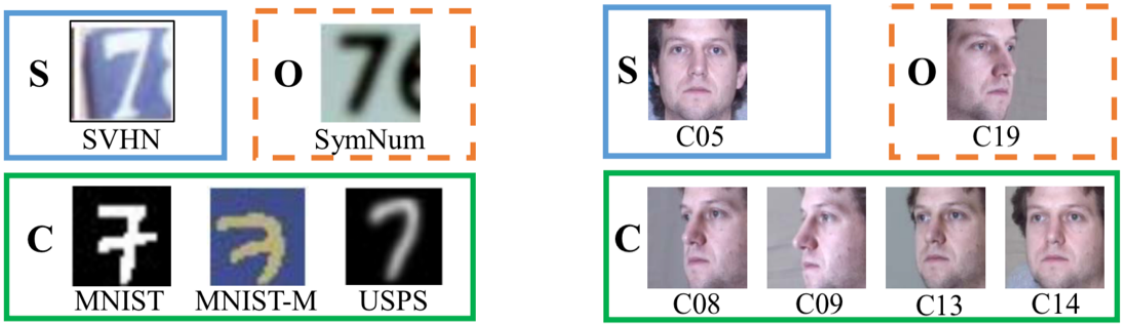
$$v_{transfer} = v_{direct} + e_{domain} \otimes v_{enhance}$$



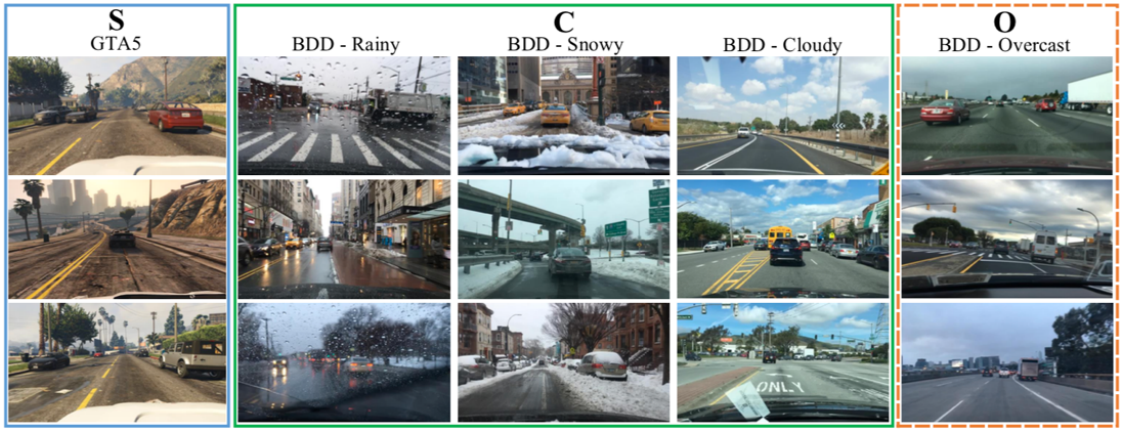

Source
Compound Targets
Open Targets
Simulation
Open World Driving Conditions
...
...
Cloudy
Rainy
Overcast
instance-wise curriculum
domain memory
Domain Disentanglement
Adaptive Knowledge Transfer
M
Continuous Adaptation

C-Digits Benchmark
Absolute Performance Gain: ~5%

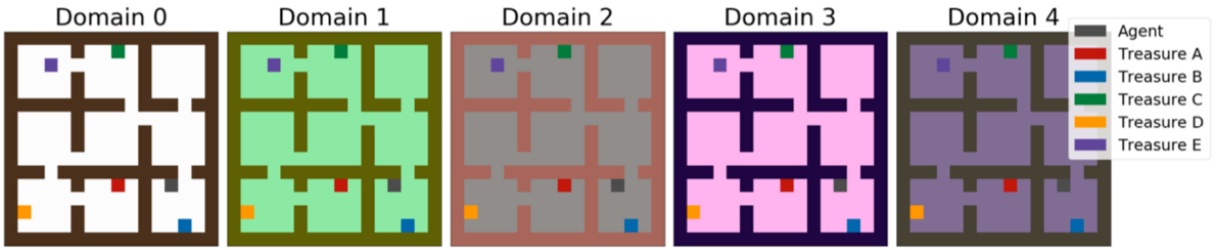
C-Faces Benchmark
Absolute Performance Gain: ~10%



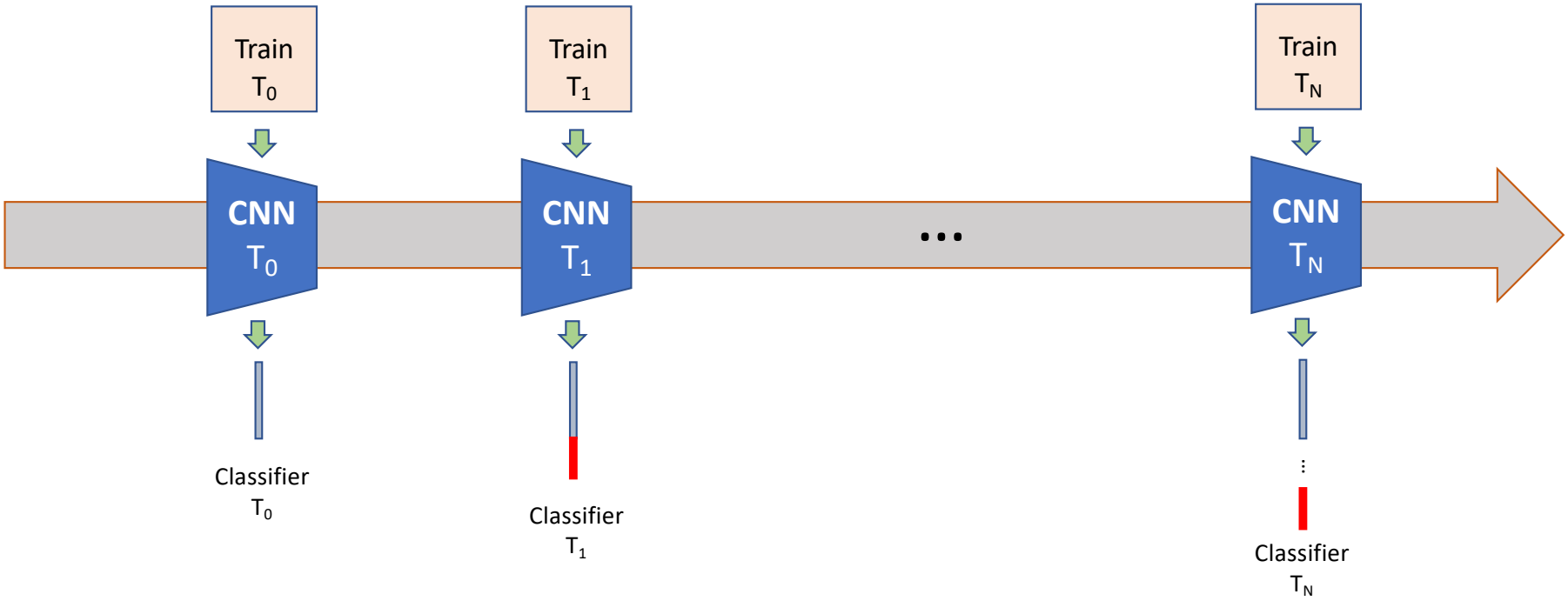
C-Driving Benchmark
Absolute Performance Gain: ~2%



C-Mazes Benchmark
Absolute Performance Gain: ~30%

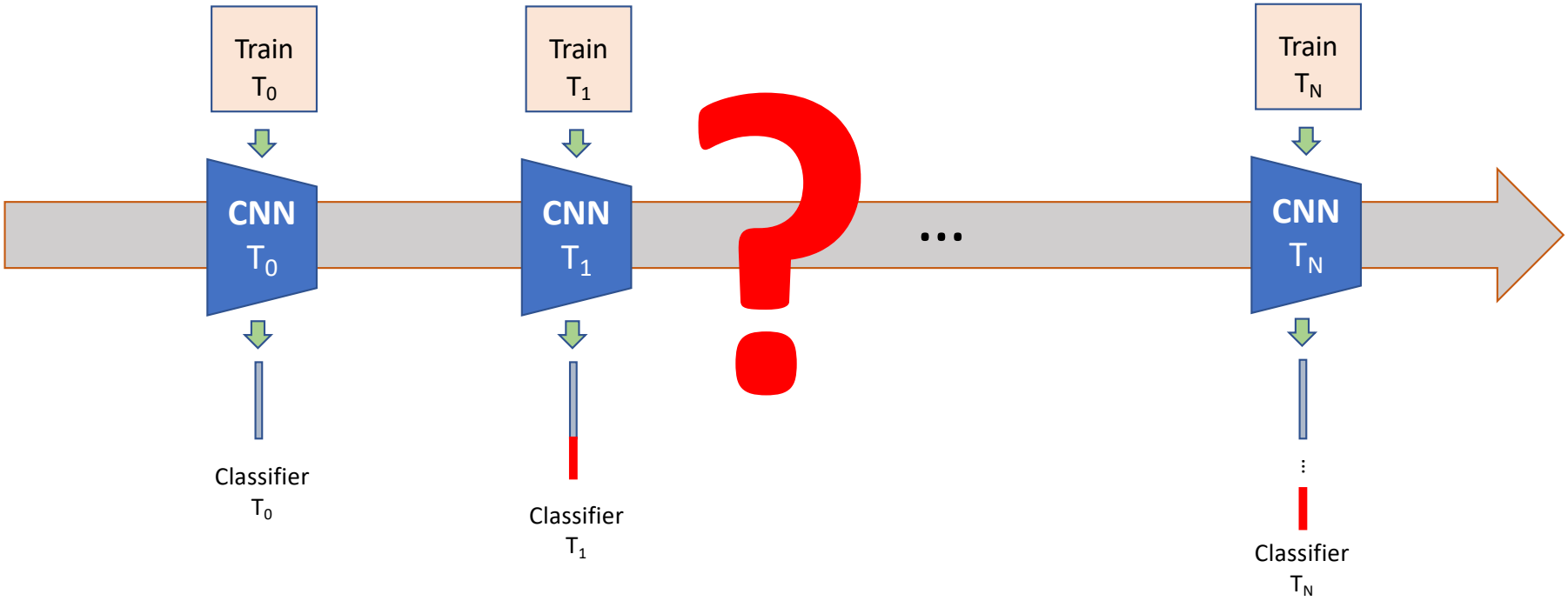


Problem continues



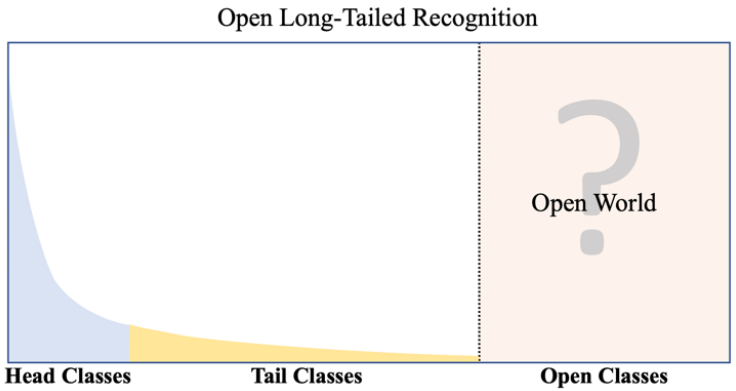
Deep learning system

Problem continues

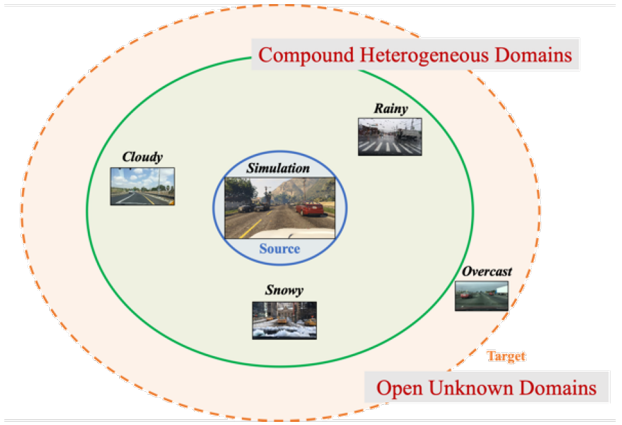


Realistic machine learning system

Thank you!



OLTR



OCDA