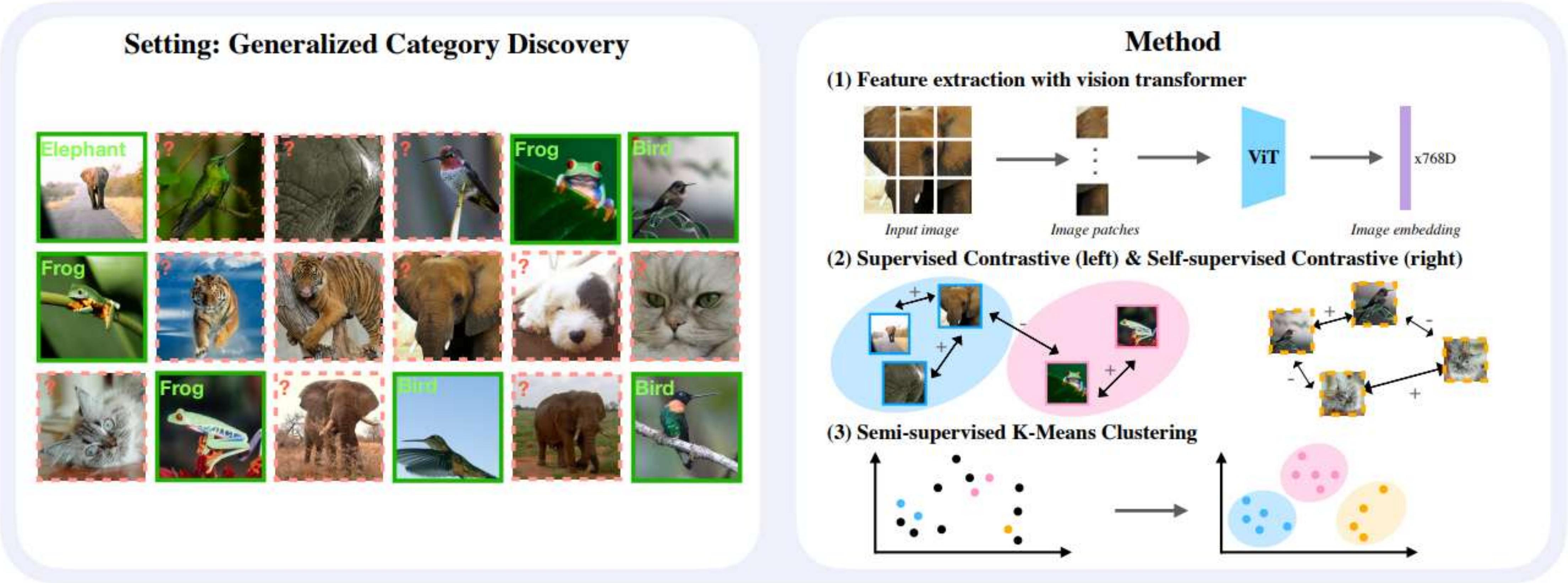
2025 ICCV

A Hidden Stumbling Block in Generalized Category Discovery: Distracted Attention

Authors:Qiyu Xu, Zhanxuan Hu, Yu Duan, Ercheng Pei, Yonghang Tai

Background



Generalized category discovery can be seen as semi-supervised learning in open scenarios, where data is divided into labeled and unlabeled parts. The labeled part contains only old class samples, while the unlabeled part contains both old and new class samples.

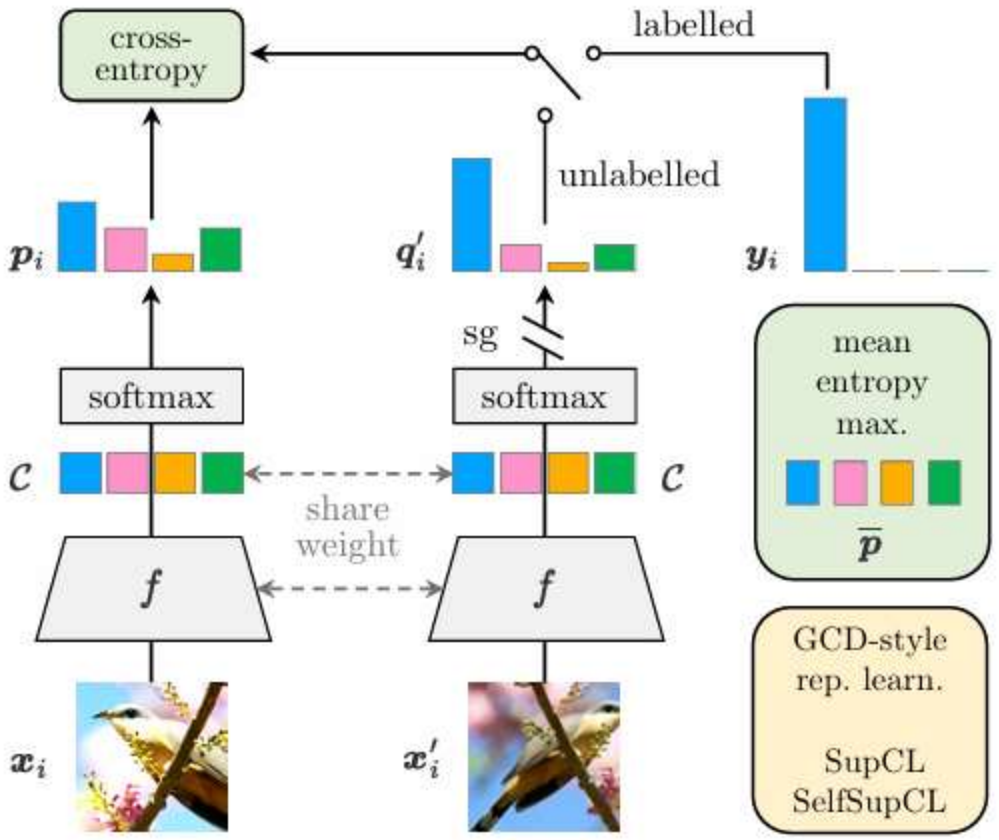
Reference： Generalized Category Discovery-CVPR 2022

Background

2023

- Open-world Semi-supervised Generalized Relation Discovery Aligned in a Real-world Setting (**EMNLP** 2023) [paper] [code]
- A Graph-Theoretic Framework for Understanding Open-World Semi-Supervised Learning (**NeurIPS** 2023) [paper] [code]
- Decompose Novel into Known: Part Concept Learning For 3D Novel Class Discovery (**NeurIPS** 2023) [paper]
- Learn to Categorize or Categorize to Learn? Self-Coding for Generalized Category Discovery (**NeurIPS** 2023) [paper] [code]
- Towards Distribution-Agnostic Generalized Category Discovery (**NeurIPS** 2023) [paper] [code]
- No Representation Rules Them All in Category Discovery (**NeurIPS** 2023) [paper] [code]
- Discover and Align Taxonomic Context Priors for Open-world Semi-Supervised Learning (**NeurIPS** 2023) [paper] [code]
- Generalized Category Discovery with Clustering Assignment Consistency (**ICONIP** 2023) [paper]
- Towards Novel Class Discovery: A Study in Novel Skin Lesions Clustering (**MICCAI** 2023) [paper]
- Novel Class Discovery for Long-tailed Recognition (**TMLR** 2023) [paper]
- Generalized Categories Discovery for Long-tailed Recognition (**ICCV Workshop** 2023) [paper]
- Boosting Novel Category Discovery Over Domains with Soft Contrastive Learning and All-in-One Classifier (**ICCV** 2023) [paper] [code]
- Parametric Information Maximization for Generalized Category Discovery (**ICCV** 2023) [paper] [code]
- MetaGCD: Learning to Continually Learn in Generalized Category Discovery (**ICCV** 2023) [paper] [code]
- Proxy Anchor-based Unsupervised Learning for Continuous Generalized Category Discovery (**ICCV** 2023) [paper] [code]
- Class-relation Knowledge Distillation for Novel Class Discovery (**ICCV** 2023) [paper]
- Incremental Generalized Category Discovery (**ICCV** 2023) [paper] [code]
- Learning Semi-supervised Gaussian Mixture Models for Generalized Category Discovery (**ICCV** 2023) [paper] [code]
- Parametric Classification for Generalized Category Discovery: A Baseline Study (**ICCV** 2023) [paper] [code]
- An Interactive Interface for Novel Class Discovery in Tabular Data (**ECML PKDD** 2023, Demo Track) [paper] [code]
- When and How Does Known Class Help Discover Unknown Ones? Provable Understandings Through Spectral Analysis (**ICML** 2023) [paper] [code]
- Open-world Semi-supervised Novel Class Discovery (**IJCAI** 2023) [paper] [code]
- ImbaGCD: Imbalanced Generalized Category Discovery (**CVPR Workshop** 2023) [paper]
- On-the-Fly Category Discovery (**CVPR** 2023) [paper] [code]
- Bootstrap Your Own Prior: Towards Distribution-Agnostic Novel Class Discovery (**CVPR** 2023) [paper] [code]
- Dynamic Conceptional Contrastive Learning for Generalized Category Discovery (**CVPR** 2023) [paper] [code]
- PromptCAL: Contrastive Affinity Learning via Auxiliary Prompts for Generalized Novel Category Discovery (**CVPR** 2023) [paper] [code]
- Modeling Inter-Class and Intra-Class Constraints in Novel Class Discovery (**CVPR** 2023) [paper] [code]
- Novel Class Discovery for 3D Point Cloud Semantic Segmentation (**CVPR** 2023) [paper] [code]
- Generalized Category Discovery with Decoupled Prototypical Network (**AAAI** 2023) [paper] [code] (DPN)
- Supervised Knowledge May Hurt Novel Class Discovery Performance (**TMLR** 2023) [paper][code]
- OpenCon: Open-world Contrastive Learning (**TMLR** 2023) [paper] [code]

Background



SimGCD combines representation learning with classifier learning and integrates self distillation algorithm to process unlabeled data. The algorithm framework is simple and effective.

Reference： Parametric Classification for Generalized Category Discovery: A Baseline Study-ICCV 2023

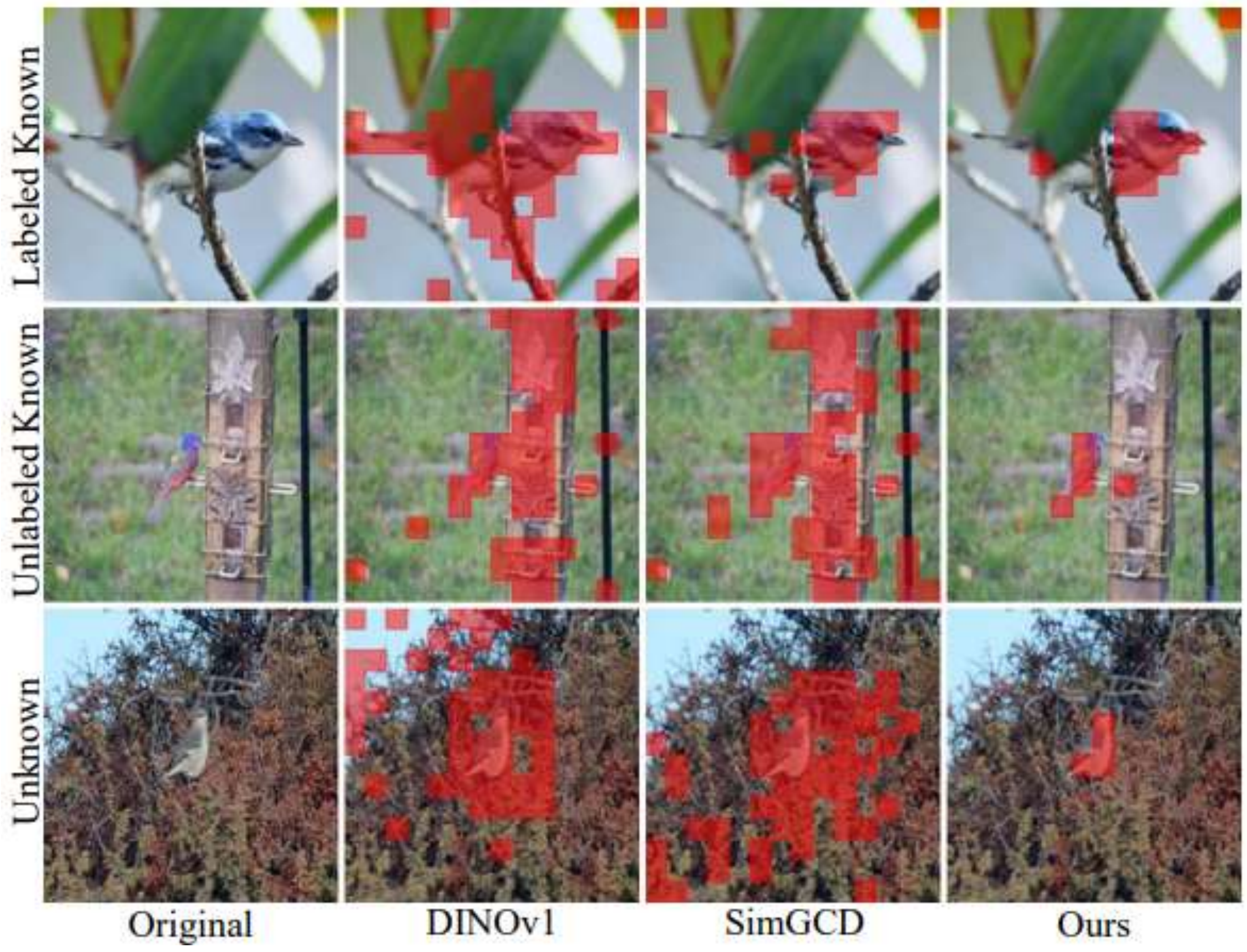
Background

2024

- Novel class discovery meets foundation models for 3D semantic segmentation (**IJCV** 2024) [paper] [code]
- Prototypical Hash Encoding for On-the-Fly Fine-Grained Category Discovery (**NeurIPS** 2024) [paper] [code]
- Happy: A Debiased Learning Framework for Continual Generalized Category Discovery (**NeurIPS** 2024) [paper] [code]
- Textual Knowledge Matters: Cross-Modality Co-Teaching for Generalized Visual Class Discovery (**ECCV** 2024) [paper] [code]
- SelEx: Self-Expertise in Fine-Grained Generalized Category Discovery (**ECCV** 2024) [paper] [code]
- Online Continuous Generalized Category Discovery (**ECCV** 2024) [paper] [code]
- PromptCCD: Learning Gaussian Mixture Prompt Pool for Continual Category Discovery (**ECCV** 2024) [paper] [code]
- Self-Cooperation Knowledge Distillation for Novel Class Discovery (**ECCV** 2024) [paper]
- Dual-level Adaptive Self-Labeling for Novel Class Discovery in Point Cloud Segmentation (**ECCV** 2024) [paper] [code]
- Contextuality Helps Representation Learning for Generalized Category Discovery (**ICIP** 2024) [paper] [code]
- NC-NCD: Novel Class Discovery for Node Classification (**CIKM** 2024) [paper]
- A Practical Approach to Novel Class Discovery in Tabular Data (**DMKD** 2024) [paper] [code]
- Novel Class Discovery for Ultra-Fine-Grained Visual Categorization (**CVPR** 2024) [paper] [code]
- Contrastive Mean-Shift Learning for Generalized Category Discovery (**CVPR** 2024) [paper] [code]
- CDAD-Net: Bridging Domain Gaps in Generalized Category Discovery (**CVPR Workshop** 2024) [paper]
- Active Generalized Category Discovery (**CVPR** 2024) [paper] [code]
- Seeing Unseen: Discover Novel Biomedical Concepts via Geometry-Constrained Probabilistic Modeling (**CVPR** 2024) [paper]
- Federated Generalized Category Discovery (**CVPR** 2024) [paper]
- Democratizing Fine-grained Visual Recognition with Large Language Models (**ICLR** 2024) [paper] [project]
- SPTNet: An Efficient Alternative Framework for Generalized Category Discovery with Spatial Prompt Tuning (**ICLR** 2024) [paper] [code]
- A Unified Knowledge Transfer Network for Generalized Category Discovery (**AAAI** 2024)
- Novel Class Discovery in Chest X-Rays via Paired Images and Text (**AAAI** 2024) [framework]
- Semantic-Guided Novel Category Discovery (**AAAI** 2024) [paper] [code]
- Adaptive Discovering and Merging for Incremental Novel Class Discovery (**AAAI** 2024) [paper]
- Debiased Novel Category Discovering and Localization (**AAAI** 2024) [paper]
- Transfer and Alignment Network for Generalized Category Discovery (**AAAI** 2024) [paper] [code]
- Guided Cluster Aggregation: A Hierarchical Approach to Generalized Category Discovery (**WACV** 2024) [paper] [code]
- AMEND: Adaptive Margin and Expanded Neighborhood for Efficient Generalized Category Discovery (**WACV** 2024) [paper] [code]
- Prediction consistency regularization for Generalized Category Discovery (**Information Fusion** 2024) [paper] [code]

The existing research methods mainly utilize other smi-supervised/unsupervised learning techniques to improve the learning of unlabeled data

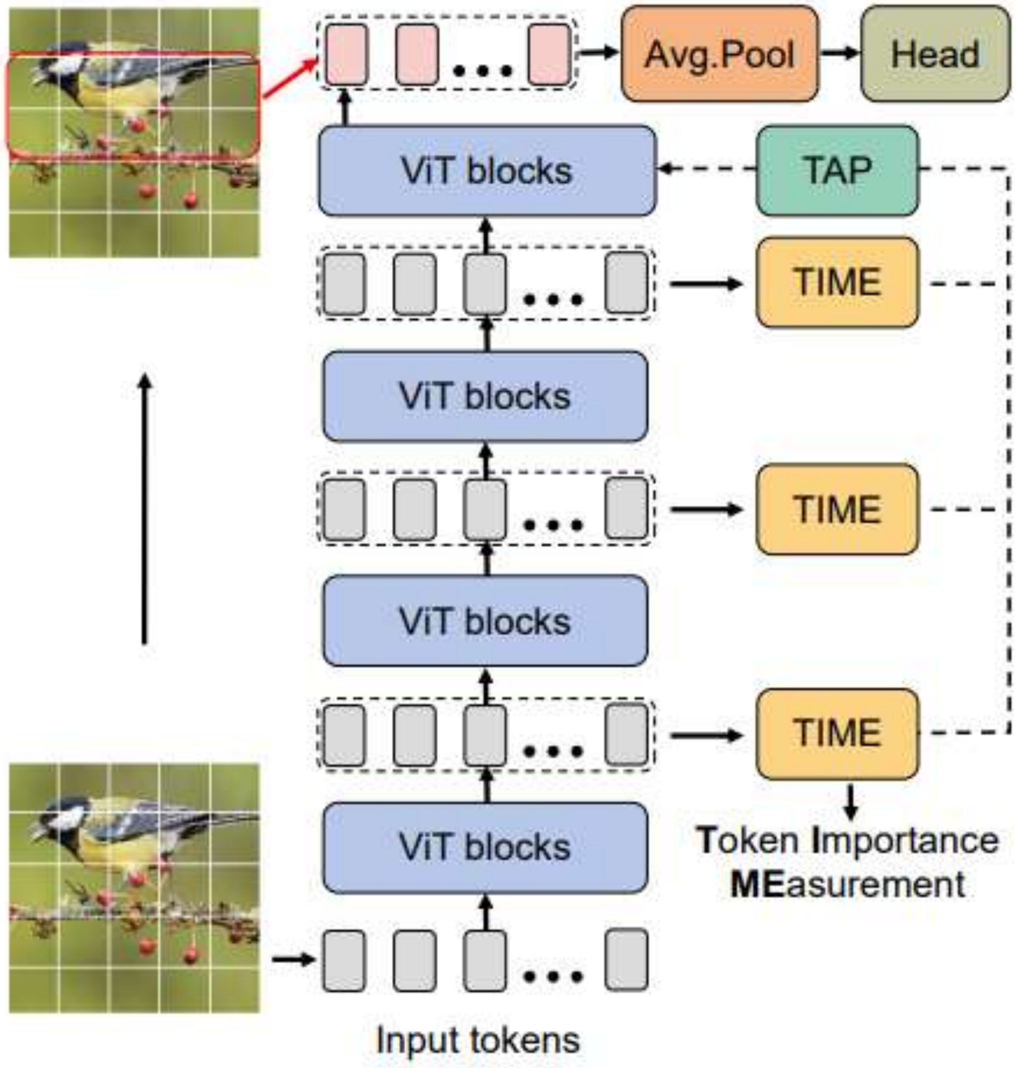
Motivation: Distracted Attention



Starting point: Why does unlabeled data have poor learning performance

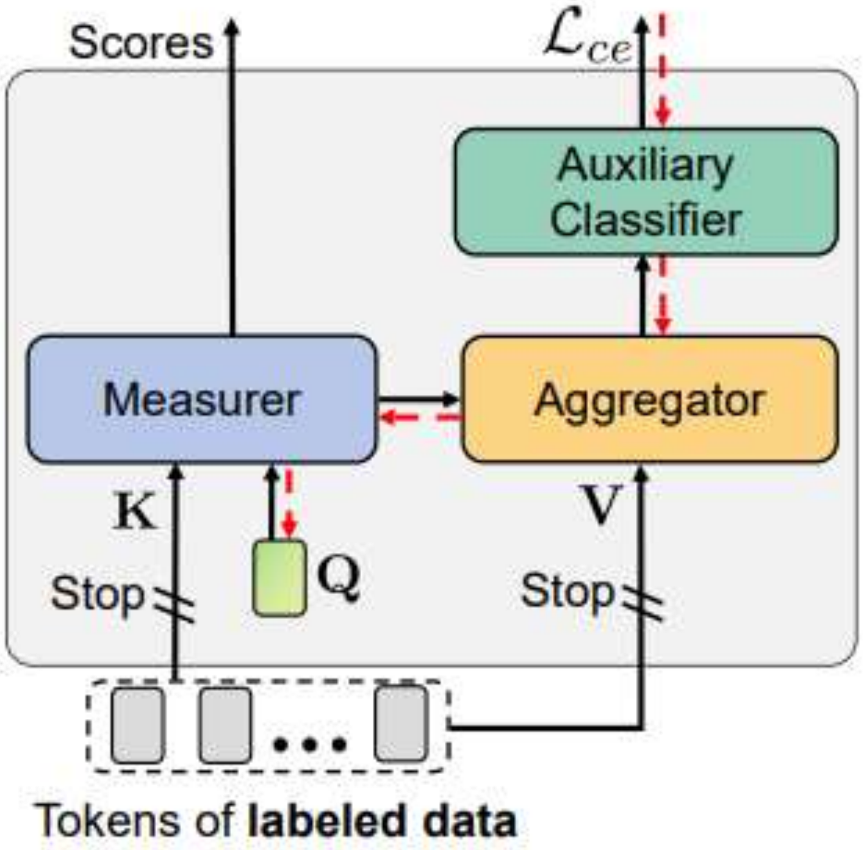
Reason: There is label guidance for labeled data, while unlabeled data may use backgrounds with small changes as classification shortcuts, leading to model attention drift.

Method



Add **TIME** module to the first 11 layers of blocks for importance measurement, generate clipping masks in **TAP** as input to the last layer of blocks, and pool the output of the last layer to generate feature input *Head*.

Method



Importance measurement：

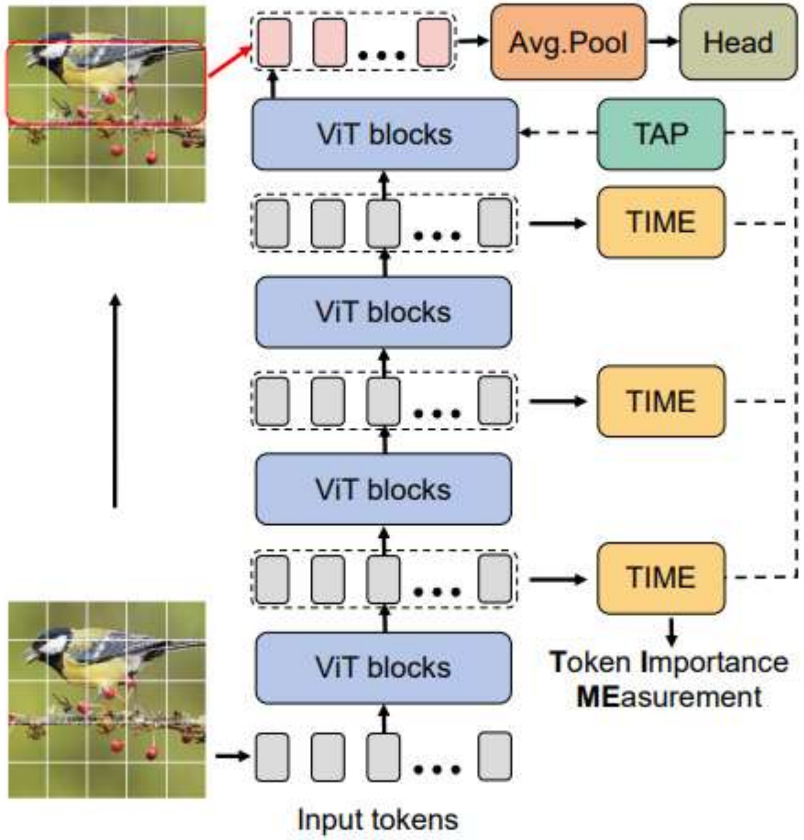
$$\mathbf{s}(\mathbf{Q}, \mathbf{K}) = \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{D}}$$

Token aggregation:

$$\mathbf{r} = \text{Softmax}(\mathbf{s})\mathbf{V}$$

TIME (Token Importance Measurement Module)

Method



The **TAP** module fuses the **TIME** outputs of the first 11 layers and uses a threshold to generate a Token Mask, which is then input into the last layer block to block the clipped Token from participating in attention calculation.

$$\mathbf{s}^m = \frac{1}{L-1} \sum_{l=1}^{L-1} \mathrm{Softmax}(\hat{\mathbf{s}}_l)$$

$$\mathbf{X}_p = \{\mathbf{x}_i \mid i = 1, 2, \ldots, t, \sum_{i=1}^{N} s_i^m \leq \tau\}$$

Experiments Results

Datasets	CUB			Stanford Cars			FGVC-Aircraft		
	All	Old	New	All	Old	New	All	Old	New
RankStats [13]	33.3	51.6	24.2	28.3	61.8	12.1	26.9	36.4	22.2
UNO+ [9]	35.1	49.0	28.1	35.5	70.5	18.6	40.3	56.4	32.2
ORCA [12]	35.3	45.6	30.2	23.5	50.1	10.7	22.0	31.8	17.1
GCD [31]	51.3	56.6	48.7	39.0	57.6	29.9	45.0	41.1	46.9
DCCL [24]	63.5	60.8	64.9	43.1	55.7	36.2	-	-	-
GPC [40]	55.4	58.2	53.1	42.8	59.2	32.8	46.3	42.5	47.9
PIM [5]	62.7	75.7	56.2	43.1	66.9	31.6	-	-	-
InfoSieve [26]	69.4	**77.9**	65.2	55.7	74.8	46.4	56.3	63.7	52.5
CMS [6]	68.2	76.5	64.0	56.9	76.1	47.6	56.0	63.4	52.3
SPTNet [33]	65.8	68.8	65.1	59.0	79.2	49.3	59.3	61.8	58.1
AptGCD [38]	**70.3**	74.3	**69.2**	62.1	79.7	53.6	**61.1**	65.2	**59.0**
MOS [23]	69.6	72.3	68.2	64.6	**80.9**	56.7	**61.1**	66.9	58.2
SimGCD [35]	60.3	65.6	57.7	53.8	71.9	45.0	54.2	59.1	51.8
SimGCD+AF	69.0	74.3	66.3	**67.0**	80.7	**60.4**	59.4	**68.1**	55.0
△	+8.7	+8.7	+8.6	+13.2	+8.8	+15.4	+5.2	+9.0	+3.2

Fine-grained Datasets

Datasets	CUB			Stanford Cars			FGVC-Aircraft		
	All	Old	New	All	Old	New	All	Old	New
CMS	67.3	75.6	63.1	53.1	73.0	43.5	54.2	63.2	49.8

CUB

Stanford Cars

FGVC-Aircraft

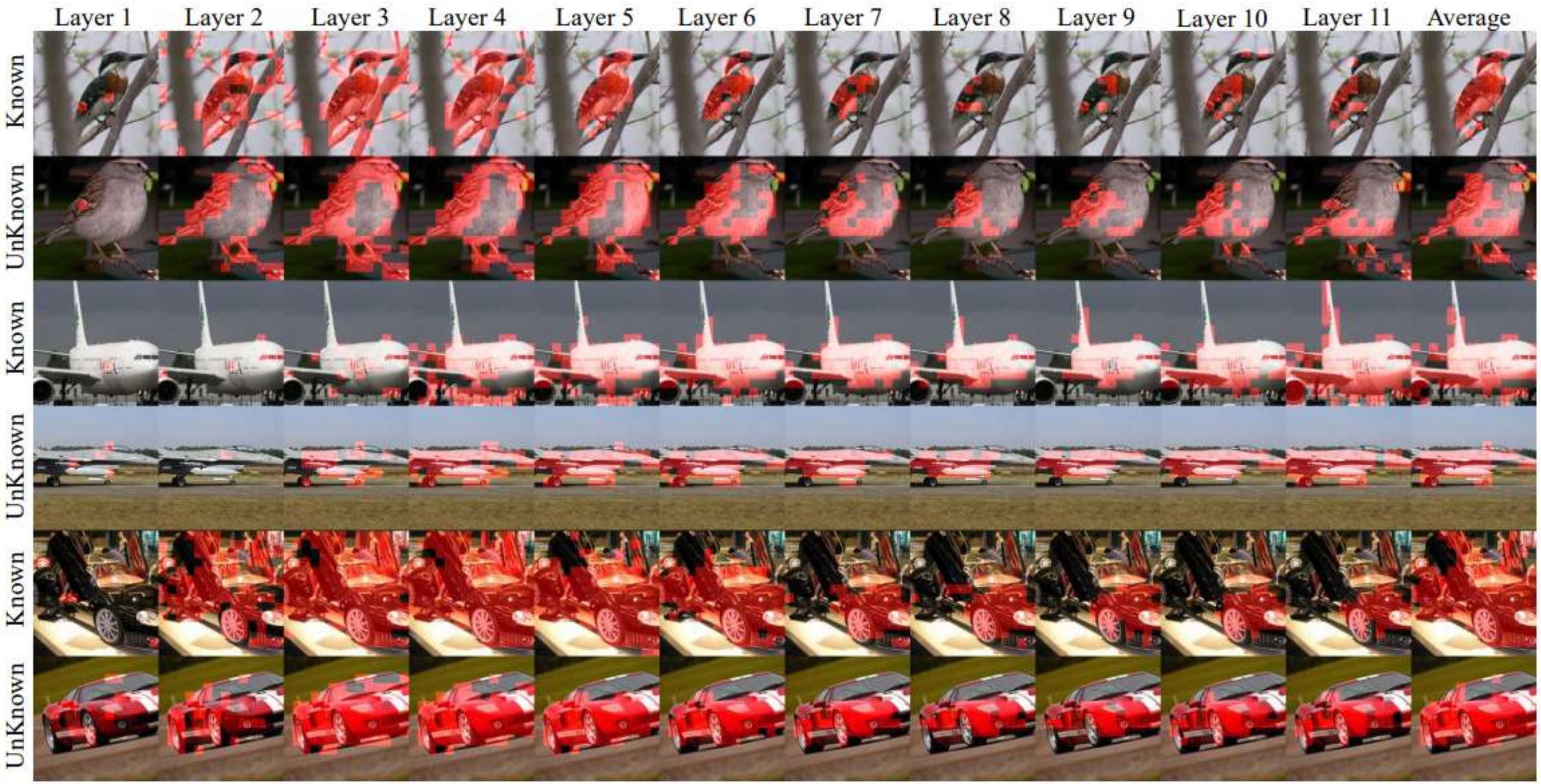
Experiments Results

Datasets	CIFAR10			CIFAR100			ImageNet-100		
	All	Old	New	All	Old	New	All	Old	New
RankStats [13]	46.8	19.2	60.5	58.2	77.6	19.3	37.1	61.6	24.8
UNO+ [9]	68.6	**98.3**	53.8	69.5	80.6	47.2	70.3	95.0	57.9
ORCA [12]	96.9	95.1	97.8	69.0	77.4	52.0	73.5	92.6	63.9
GCD [31]	91.5	97.9	88.2	73.0	76.2	66.5	74.1	89.8	66.3
DCCL [24]	96.3	96.5	96.9	75.3	76.8	70.2	80.5	90.5	76.2
GPC [40]	92.2	98.2	89.1	77.9	85.0	63.0	76.9	94.3	71.0
PIM [5]	94.7	97.4	93.3	78.3	84.2	66.5	83.1	95.3	77.0
InfoSieve [26]	94.8	97.7	93.4	78.3	82.2	70.5	80.5	93.8	73.8
CMS [6]	-	-	-	82.3	**85.7**	75.5	84.7	**95.6**	79.2
SPTNet [33]	97.3	95.0	98.6	81.3	84.3	75.6	85.4	93.2	81.4
AptGCD [38]	97.3	95.8	98.7	**82.8**	81.8	**85.5**	**87.8**	95.4	**84.3**
SimGCD [35]	97.1	95.1	98.1	80.1	81.2	77.8	83.0	93.1	77.9
SimGCD+AF	**97.8**	95.9	**98.8**	82.2	85.0	76.5	85.4	94.6	80.8
△	+0.7	+0.8	+0.7	+2.1	+3.8	-1.3	+2.4	+1.5	+2.9

CIFAR10 CIFAR100 ImageNet-100

Due to the good performance achieved by SimGCD on the general dataset, the model is less affected by background interference.

Experiments Results



The importance of multi-scale pruning

Experiments Results

Datasets	CUB			Stanford Cars			FGVC-Aircraft		
	All	Old	New	All	Old	New	All	Old	New
SimGCD	60.1	69.7	55.4	55.7	73.3	47.1	53.7	64.8	48.2
+AF(M-TAP)	66.8	73.1	63.6	63.2	79.9	55.1	57.4	65.7	53.3
+AF(S-TAP)	69.0	74.3	66.3	67.0	80.7	60.4	59.4	68.1	55.0

Single view pruning can be seen as an alternative form of data augmentation

THANKS