岩手県立大学
ソフトウェア情報学部
Faculty of Software and Information Science

CORE-ReID V2: Advancing the Domain Adaptation for Object Re-Identification with Optimized Training and Ensemble Fusion

CORE-ReID V2: 最適化されたトレーニングとアンサンブル融合によるオブジェクト再識別のドメイン適応の進化

Department of Software and Information Science
Supervisor: Prof. Prima Oky Dicky Ardiansyah
PhD candidate: Nguyen QuocTrinh (2362023201)
g236v201@s.iwate-pu.ac.jp

2025/07/24

Agenda

1. Research Background
2. Related Work
3. Research Aim
4. Methodology
5. Results
6. Conclusion
7. Future Work
8. References

Research Background

Needs / Issues	Motivations	Solution
Tracking individuals across multiple camera views presents challenges that traditional tracking algorithms often fail to address.	Addressing problems related to • **Security** (advanced surveillance system) • **Behavior Analysis** (behavior pattern, emotion recognition) • **Human Flow Analysis** (crowd management, simulation) • **Origin-Destination (OD) Survey** (tracking and analyzing movement)	Person Re-Identification

We also extend the method's scope to Vehicle Re-Identification.

Security

Crime Prevention CCTV (UK)

Source: [Calipsa]

Crime Prevention CCTV using Person Re-Id (China)

Source: [Financial Times]

Tokyo to Install 22,000 Security Cameras on Metro in Advance of 2020 Olympics

East Japan Railway Co., or JR East, plans to install about 22,000 security cameras as part of efforts to increase public safety and security before the 2020 Olympics

By Jessica Davis | Mar 12, 2019

East Japan Railway Co., or JR East, has **announced plans** to increase the number of security cameras at stations in and around Tokyo and set up a department to monitor the cameras 24/7. The cameras are part of the company's plan to increase public safety and security in the led up to the 2020 Olympics, which will be held in Tokyo.

According to reports, by the time the Olympics open next July, about 22,000 security cameras will be present near JR East ticket gates and on platforms at about 1,200

Source: [Security Today]

Research Background

Human Flow Analysis

Human Flow Analysis at Morioka City (2023~)

Source: https://morioka-machidukuri.jp/

Human Flow Analysis at Kochi City from (2024~)

Source: https://prtimes.jp/main/html/rd/p/000000003.000145373.html

Iwate Prefectural University

Research Background

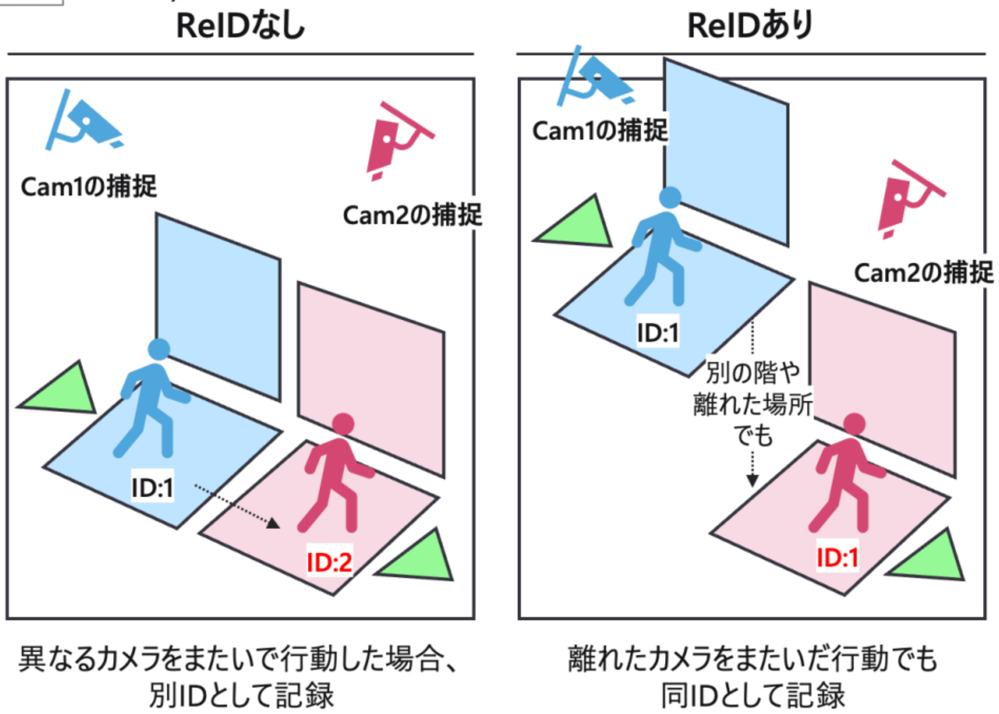
Human Flow Analysis

△駅構内に設置されたカメラ①

△駅構内に設置されたカメラ②

Human Flow Analysis Inside the JR Kitakami Station (2025/04~)

Source: https://denkikogyo.co.jp/11740/



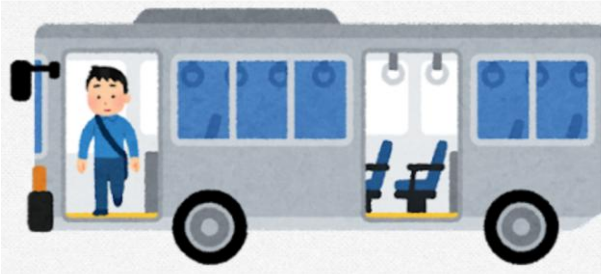
BehaveEye® and ReID Technology by Cybercore (2025/04~)

Source: https://cybercore.co.jp/news_jp/2025/2101/

Research Background

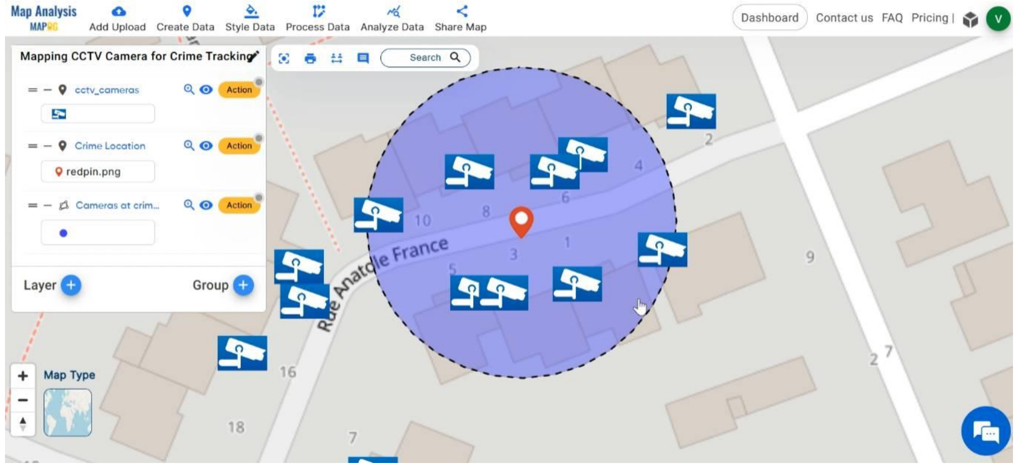
Origin-Destination Survey

Get on the bus/train



Get off the bus/train

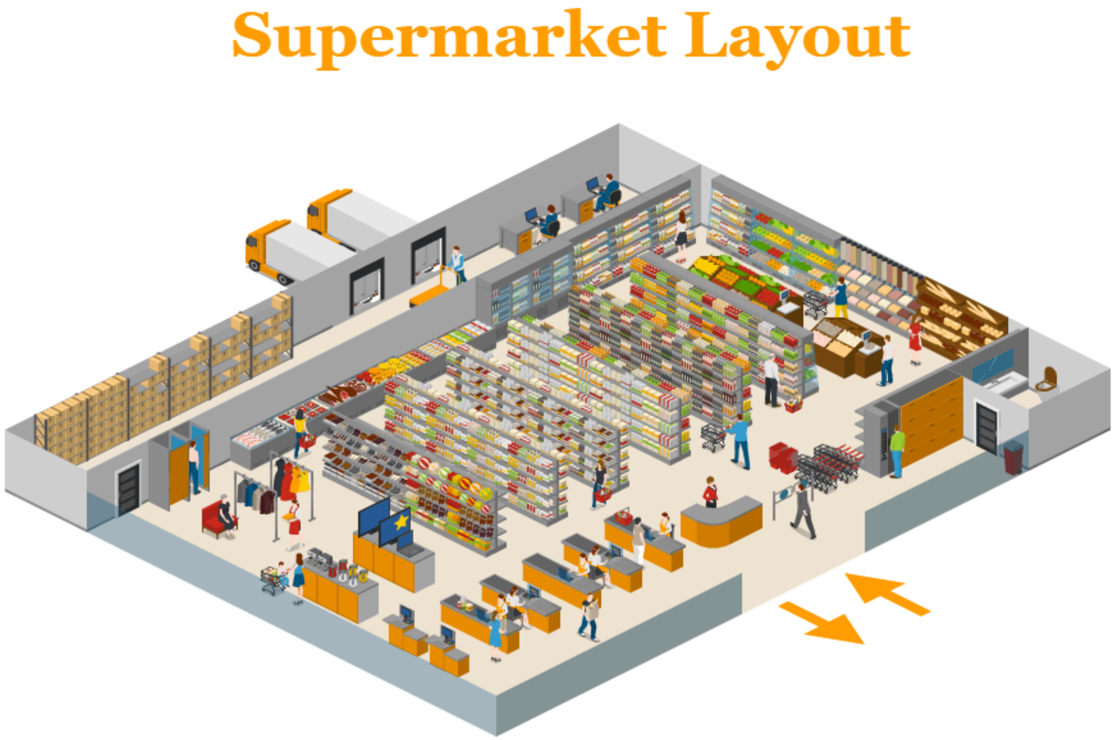
Integration with edge devices



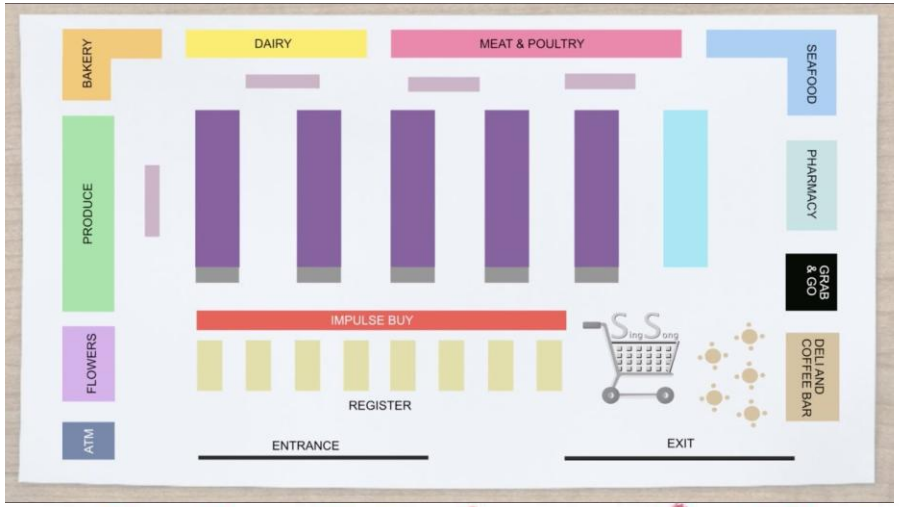
Analysis within an area

Research Background

Behavior Analysis



Shopping behavior



Iwate Prefectural University

Research Background

Person Re-Identification

Camera

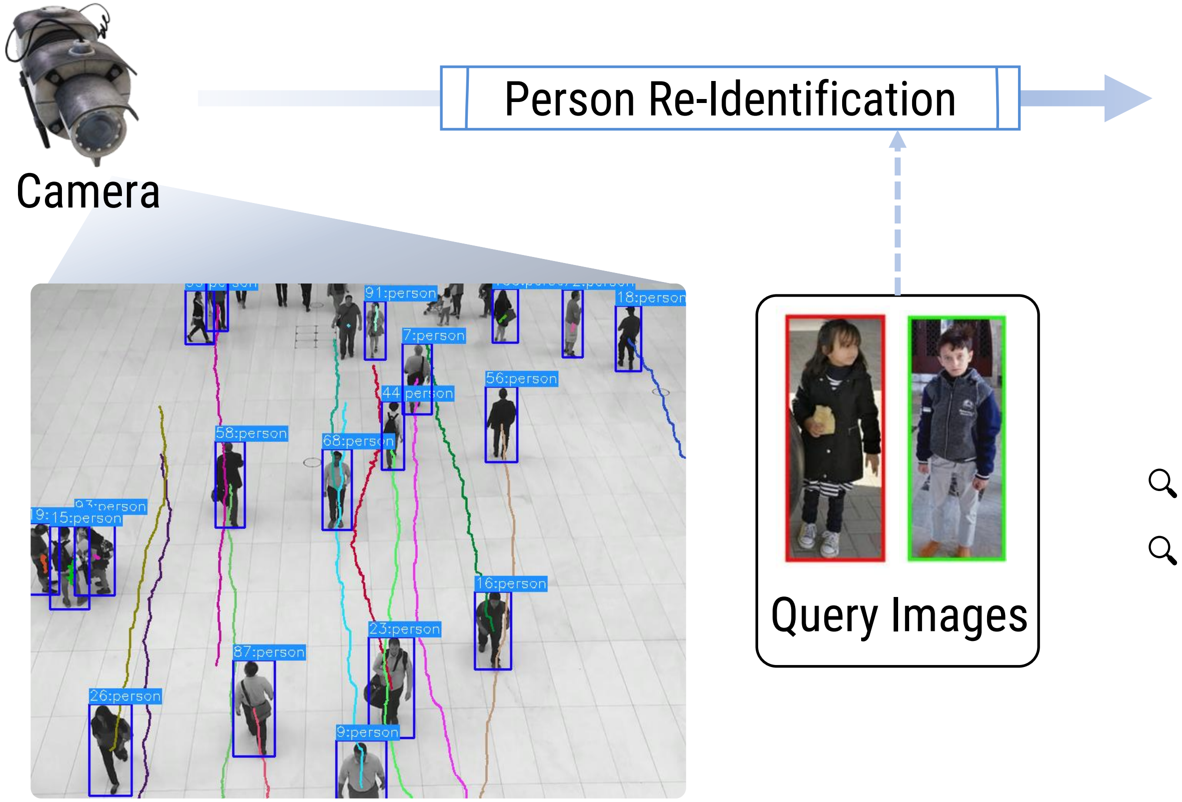
Search

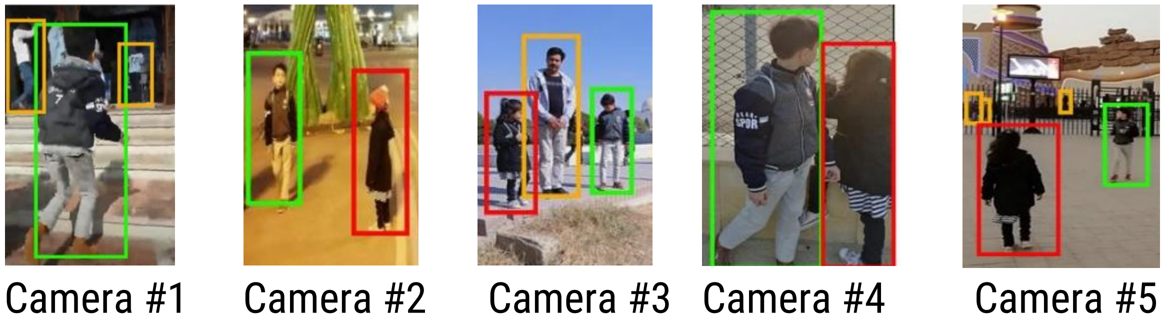
Query Image

Image Gallery

Person Re-Identification

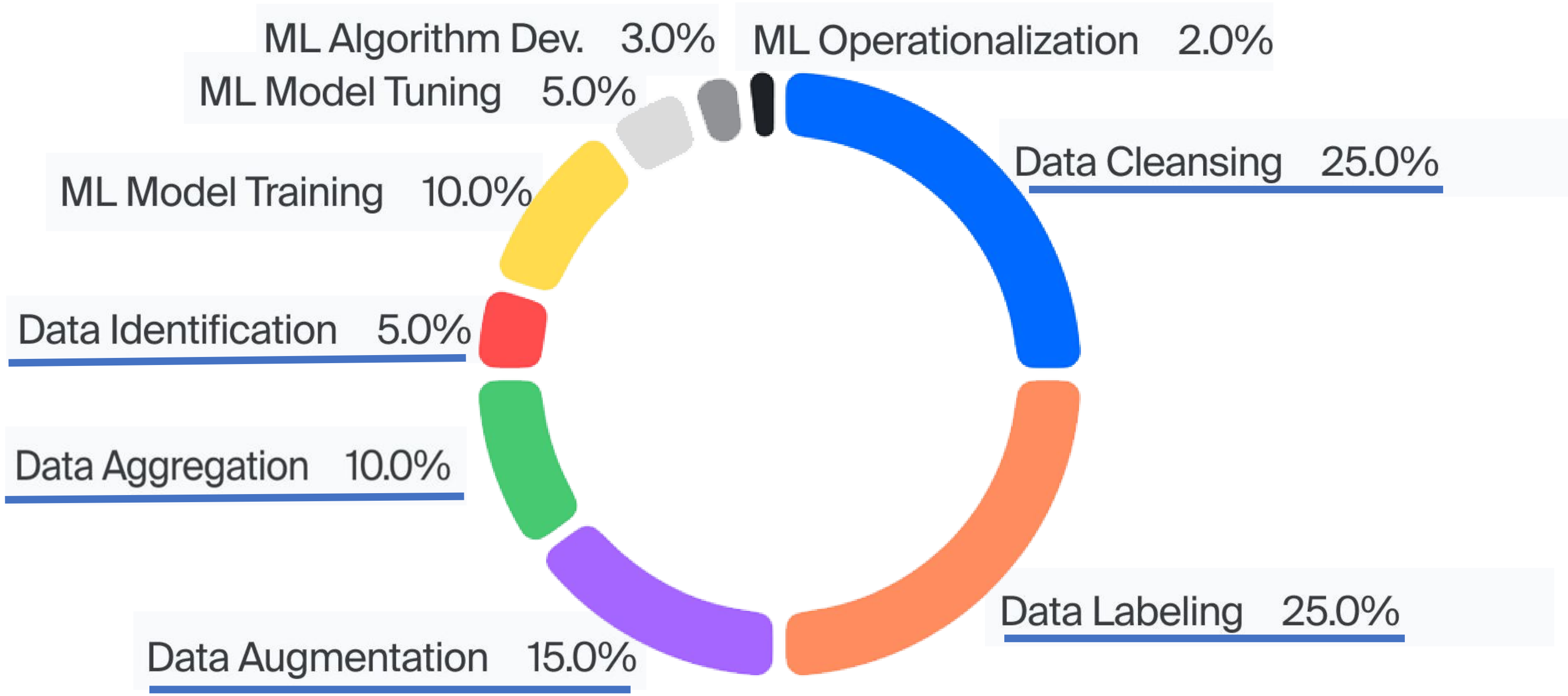
Person Re-Identification (ReID) is a computer vision task that focuses on identifying and matching individuals across non-overlapping camera views distributed at distinct locations.





- Identify, track and link individuals across multiple cameras.
- Facilitating information exchange between cameras to construct a complete trajectory of an individual's movement.

Machine Learning Project Tasks

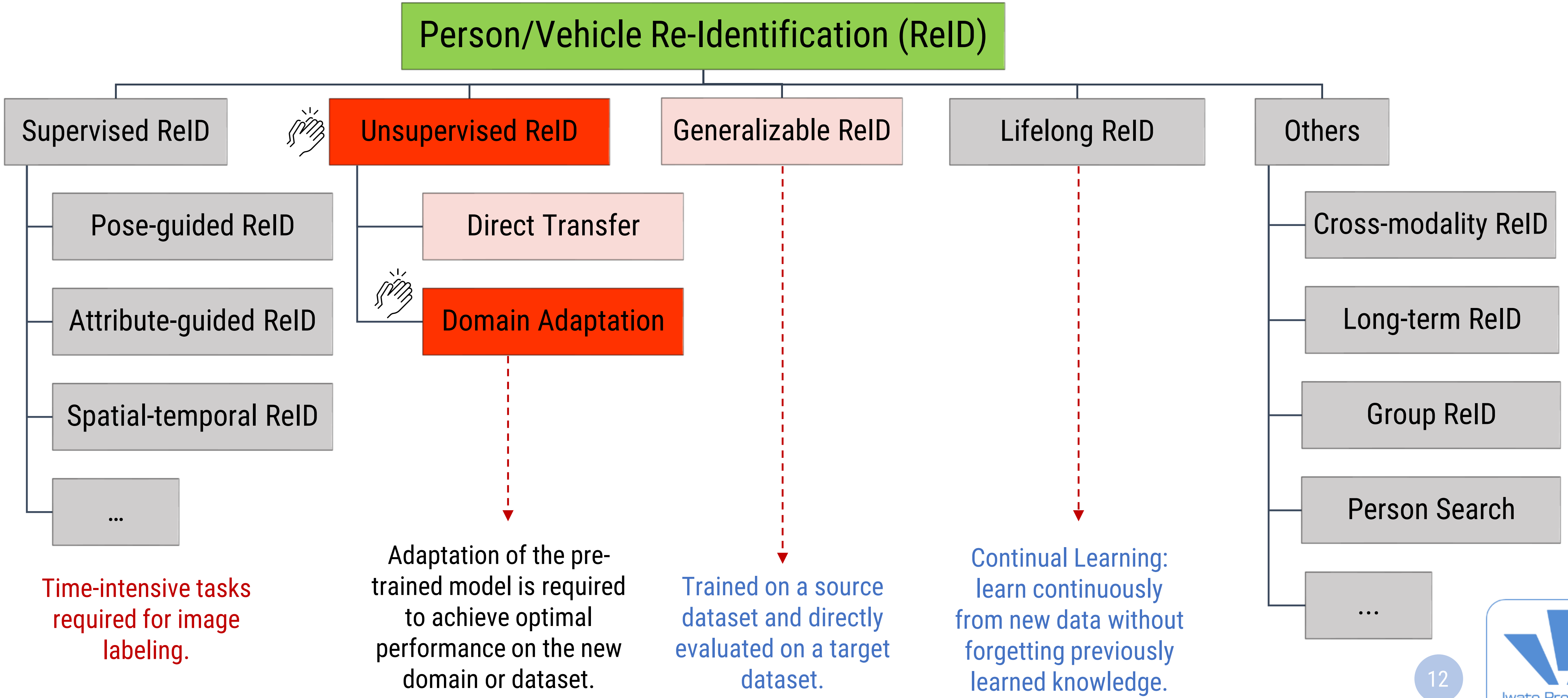


Approximately 80% of the total time is dedicated to data gathering and preprocessing, which are crucial steps for ensuring the success and accuracy of the model.

Source: Cognilityca (Percentage of Time Allocated to Machine Learning Project Tasks)

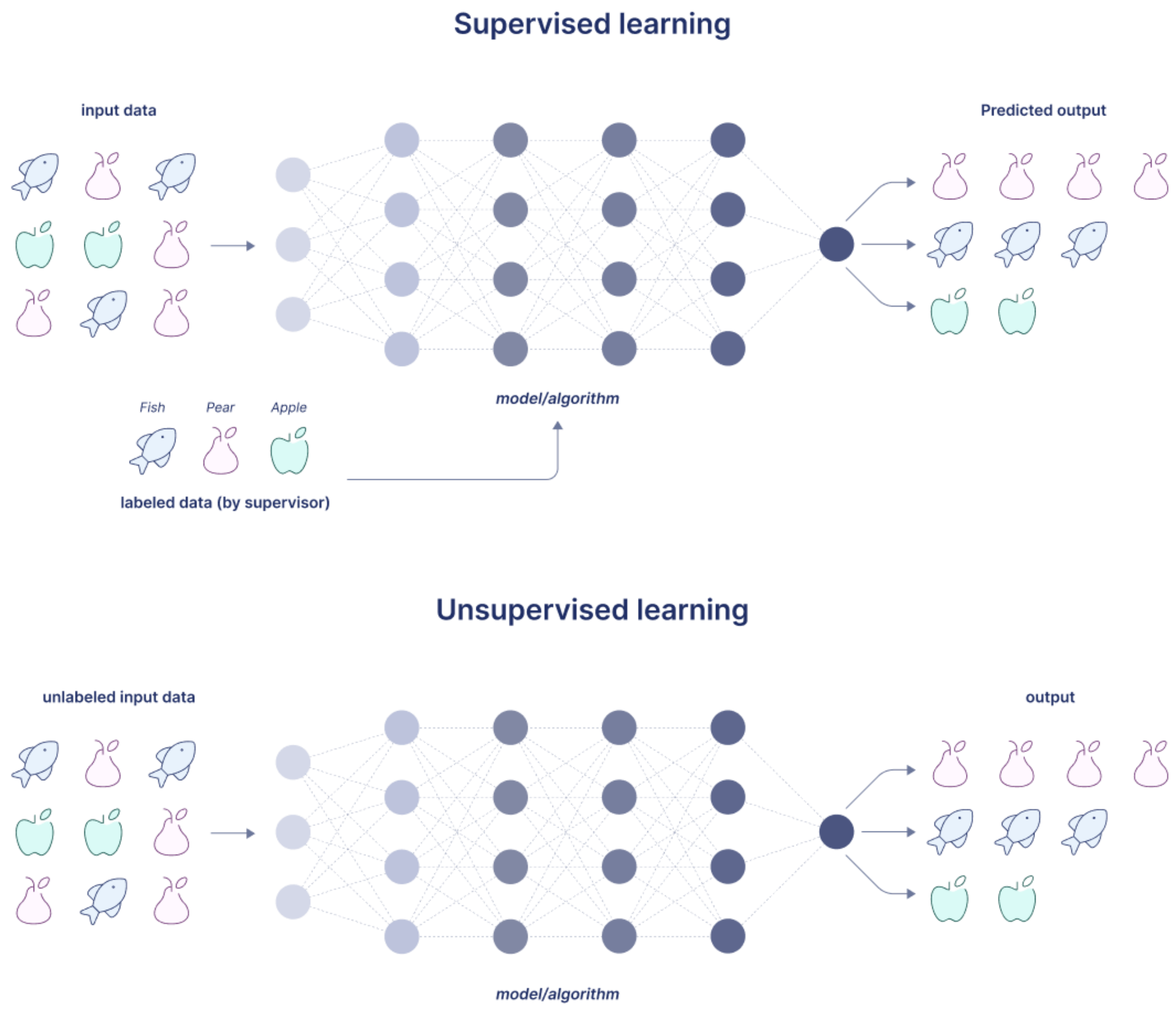
Related Work:: Person/Vehicle Re-Identification

Various Person/Vehicle Re-Identification Methods



Related Work:: Unsupervised Domain Adaption

Supervised vs Unsupervised Learning



Source: [Supervised vs. Unsupervised Learning: Key Differences]

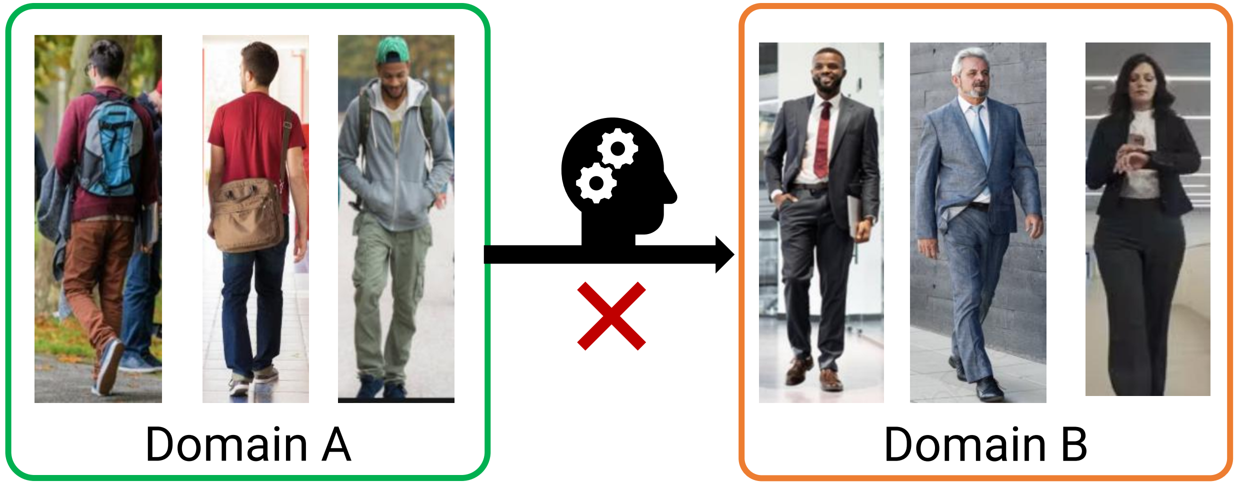
Supervised learning dominates in terms of accuracy and robustness but is resource-intensive and lacks scalability.

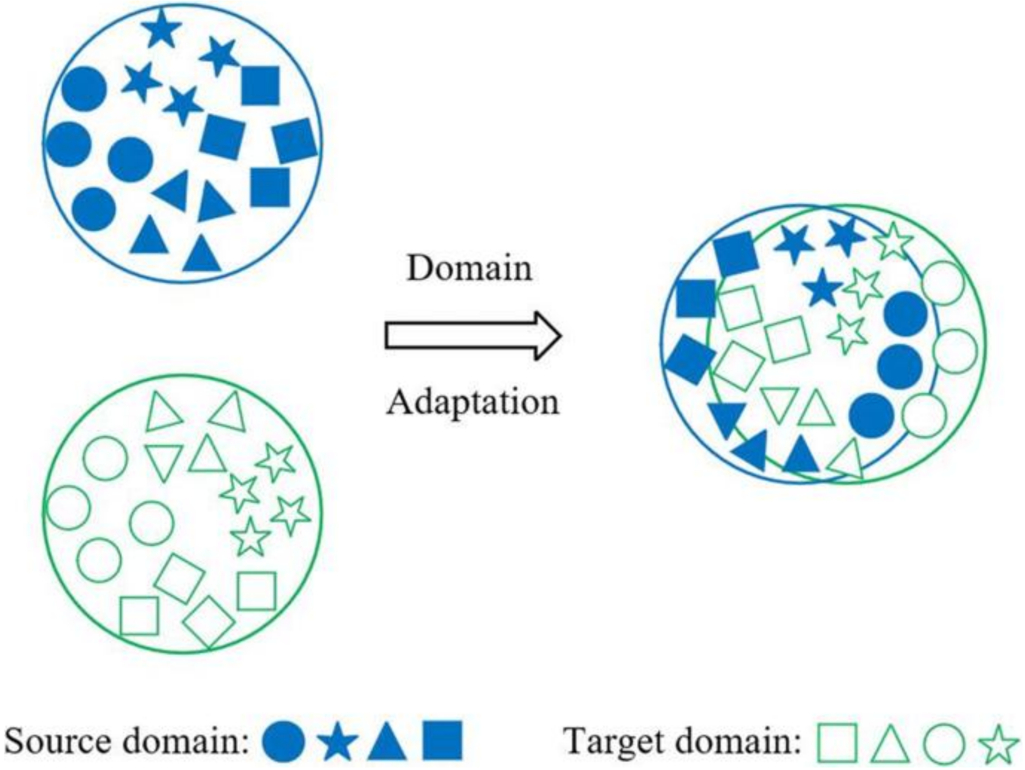
- Struggles with domain adaptation because it relies on labeled data from the target domain.
- Requires additional labeling efforts when adapting to a new domain.

Unsupervised learning, while less accurate, is more flexible, scalable, and better suited for real-world applications where data annotation is impractical.

- Works in unsupervised domain adaptation (UDA) settings by transferring knowledge from labeled source domains to unlabeled target domains.
- Often involves techniques like feature alignment, adversarial learning, and style transfer.

Cross-Domain Adaption



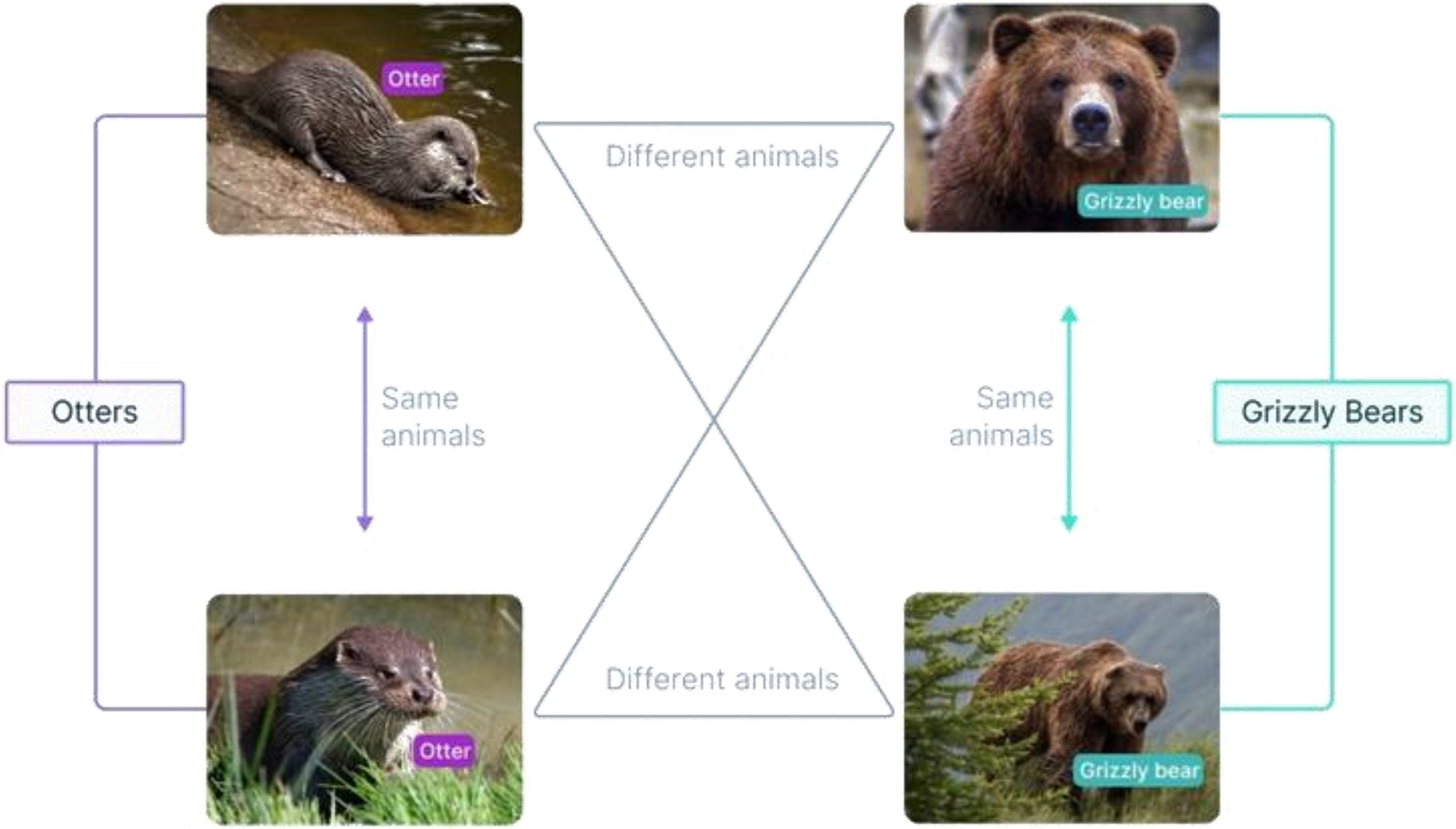


Re-identification (Re-ID) algorithms often struggle to generalize effectively across different domains.

Cross-domain adaptation for person re-identification aims to bridge the performance gap between two distinct domains.

Effective Model Training: Contrastive learning

Contrastive learning extracts meaningful representations by distinguishing between positive and negative instance pairs.

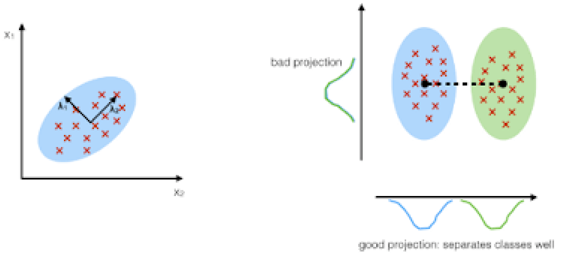


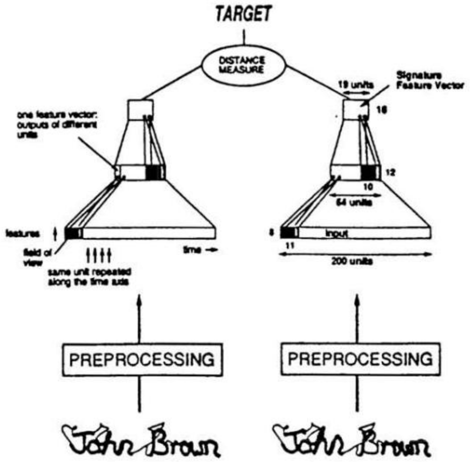
Contrastive learning (Early Foundations)

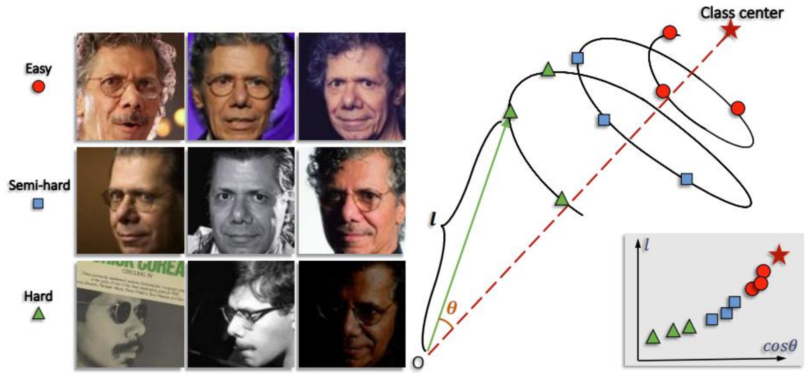
Dimensionality Reduction & Distance Metrics (1990s-2000s)	Siamese Networks (1993, Bromley et al.) [3]	Contrastive Loss (2005, Chopra et al.) [4]
Principal Comp. Analysis (PCA) [1] **Linear Discriminant Analysis (LDA)** [2].	**Siamese Networks** for signature verification	Sumit Chopra, Raia Hadsell, and **Yann LeCun** published "Learning a Similarity Metric Discriminatively, with Application to **Face Verification**" in 2005

PCA: component axes that maximize the variance

LDA: maximizing the component axes for class-separation







Related Work:: Contrastive Learning

Contrastive learning (Deep Learning Era 2010s)

Deep Siamese Networks for Face Verification (2015, FaceNet by Google) [5]

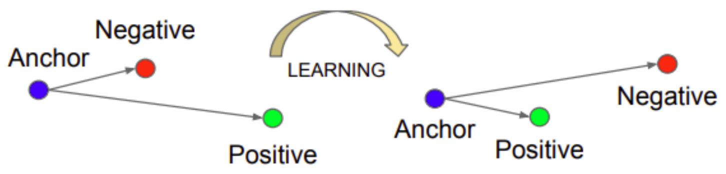
Supervised Contrastive Learning (2017-2019) [6]

Self-Supervised Contrastive Learning (2020, SimCLR & MoCo)

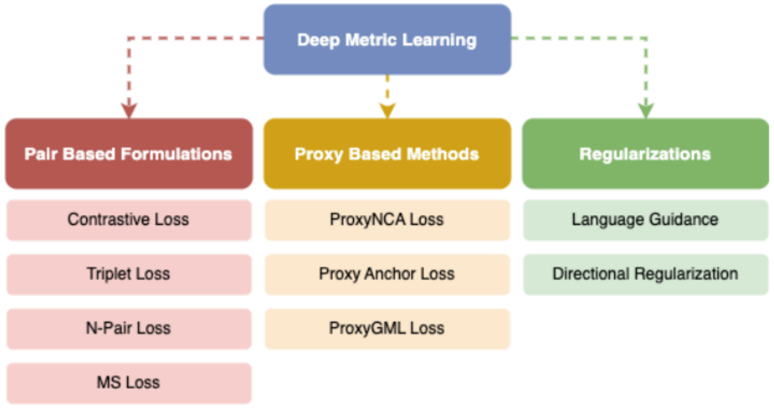
The **FaceNet** model (Schroff et al., 2015) leveraged **triplet loss**, a more advanced contrastive learning loss function, for face verification.

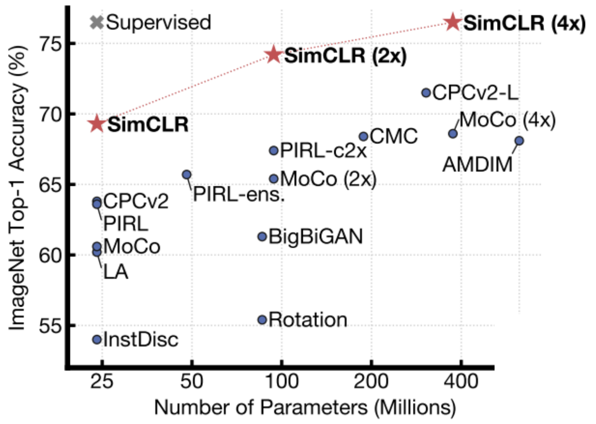
Deep Metric Learning and **Supervised Contrastive Learning** extended these ideas.

SimCLR (Chen et al., 2020) [7] and **MoCo** (He et al., 2020) [8] applied contrastive learning in Sefl-supervised tasks.



The Triplet Loss minimizes the distance between an anchor and a positive

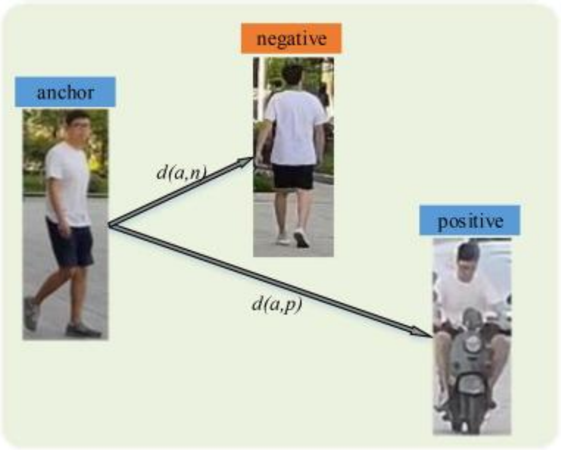


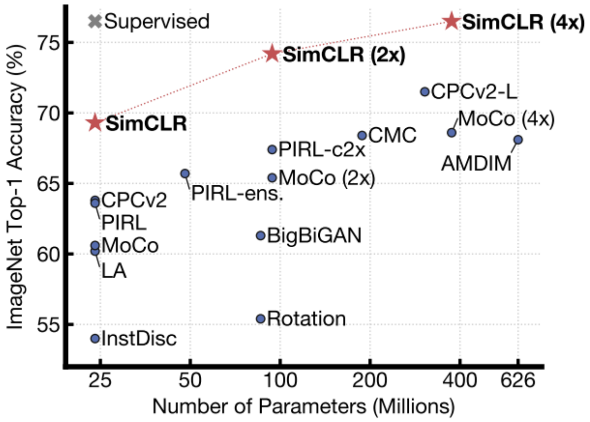


Related Work:: Contrastive Learning

Summary of related “contrastive learning” work

Early Works (2016-2018)	Modern Contrastive Learning in ReID (2019-Present)	This study: • Contrastive Learning • Ensemble (local & global features) • ReID (person and vehicle)
• Triplet Loss-based Approaches [5] • Siamese Networks for ReID [9]	• Contrastive Learning for Domain Adaptation [10] • Self-Supervised ReID [7,8]	Contrastive learning can be leveraged to enhance feature representation for ReID

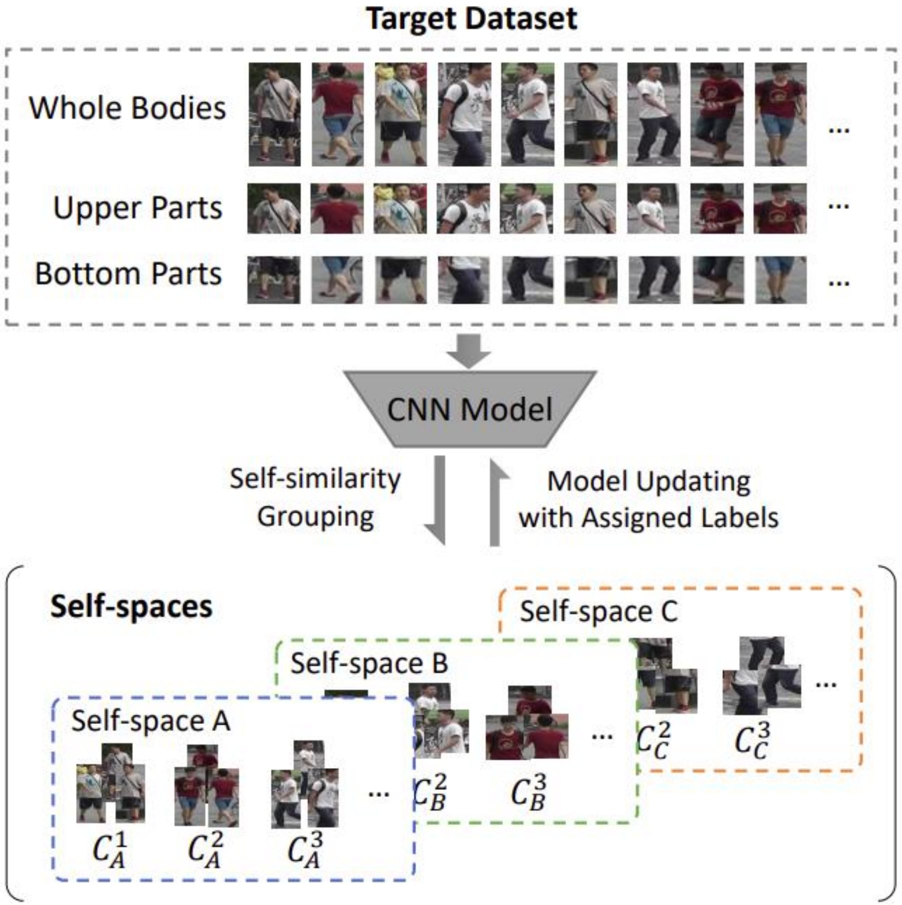




We are here

Self-similarity Grouping (SSG) [12]

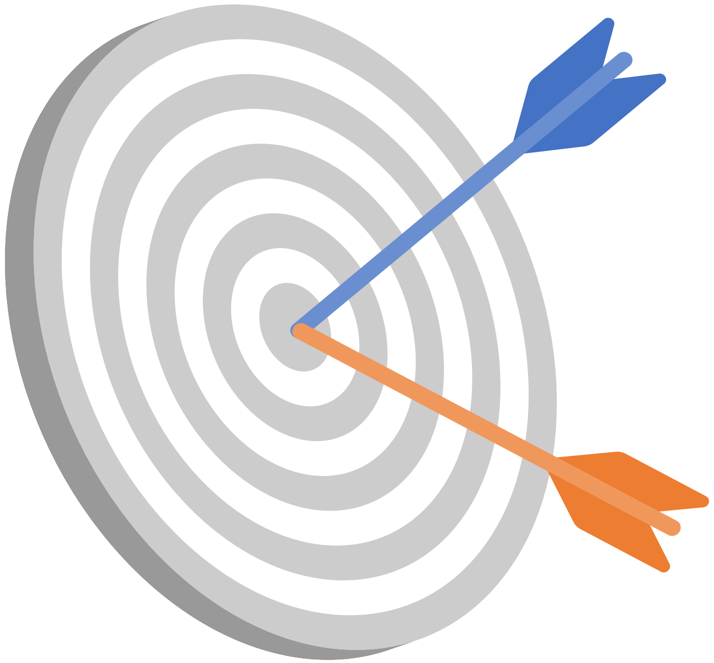
SSG explores the use of global and local features of the Unsupervised Domain Adaptation (UDA) in Person ReID.



① SSG uses a single network for feature extraction in clustering, which is susceptible to the generation of numerous noisy pseudo-labels.

② SSG performs clustering based on global and local features independently, resulting in unlabeled samples acquiring multiple different pseudo-labels, leading to ambiguity in identity classification during training.

Research Aim



Aim ①

Tackle the Unsupervised Domain Adaptation (UDA) problem in Person ReID

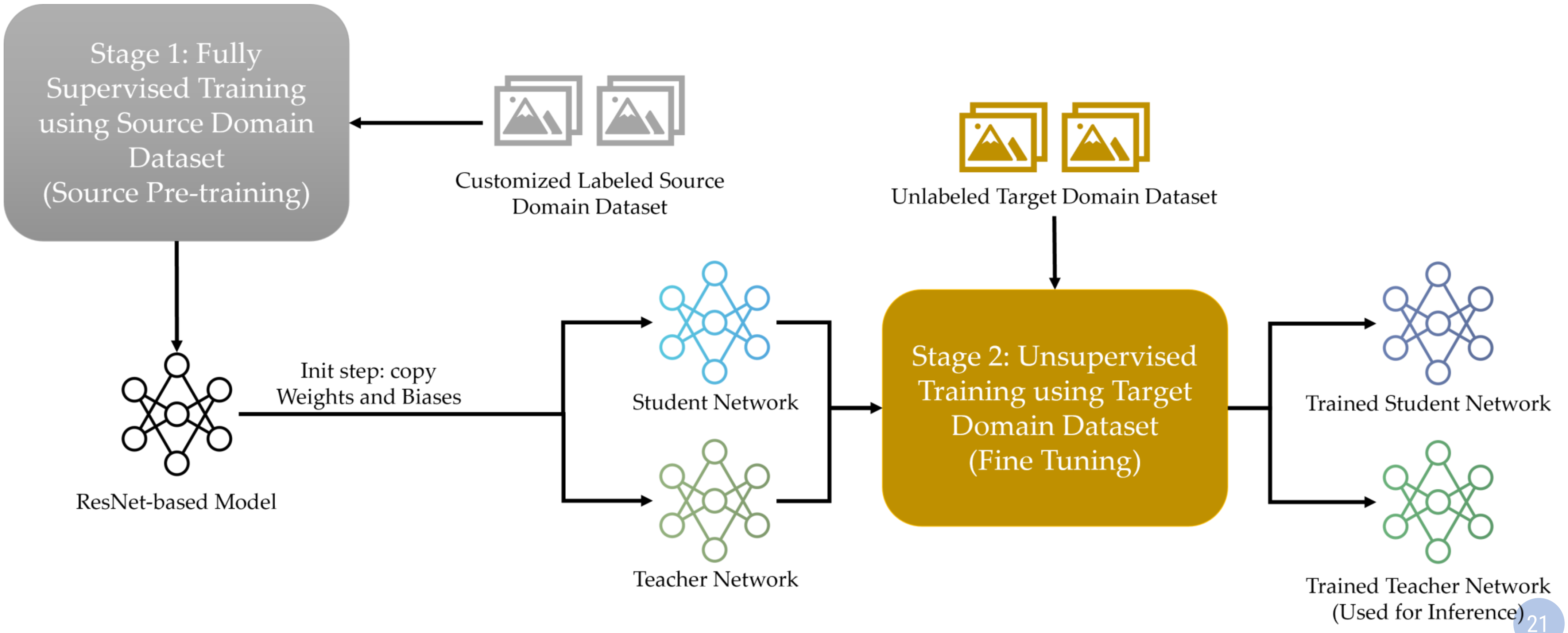
Aim ②

Extend to Vehicle ReID, and further Object ReID

Iwate Prefectural University

Proposed Methods

There are two stages: pre-training (stage 1) the model on the source domain in a fully supervised manner and fine-tuning (stage 2) the model on the target domain using an unsupervised learning approach.



Proposed Methods:: CORE-ReID V1

Needs	Related work	CORE-ReID V1
Pre-training the model on the source domain in a fully supervised manner. The data is limited; we want to use the data effectively	CAMStyle (CycleGAN) from the supervised learning task.	Use the camera-aware style transfer to improve the training dataset.
Global and local features consideration	SSG [12]	Ensemble Fusion
Control the inter-channel relationships of features to guide the model's attention to meaningful structures within the input image	Attention Map, Convolutional Block Attention Module (CBAM) [14]	Efficient Channel Attention Block (ECAB)
Extract information from input images effectively.	FlipReID [15] used horizontally flipped counterpart in supervised learning task.	Bidirectional Mean Feature Normalization (BMFN)

Proposed Methods:: CORE-ReID V2

CORE-ReID V1's Limitations	CORE-ReID V2
Limited application domain: only Person ReID	Expanded application scope: supports Person, Vehicle, and further Object ReID
Synthetic-data generation challenge: ineffective when the number of cameras was unspecified	Advanced synthetic data generation: handles unknown camera setups using both camera-aware style transfer and domain-aware style transfer
Inefficient data augmentation: The Random Grayscale Patch Replacement technique only operated locally	Improved data augmentation: combines local and global grayscale changes for stronger generalization.

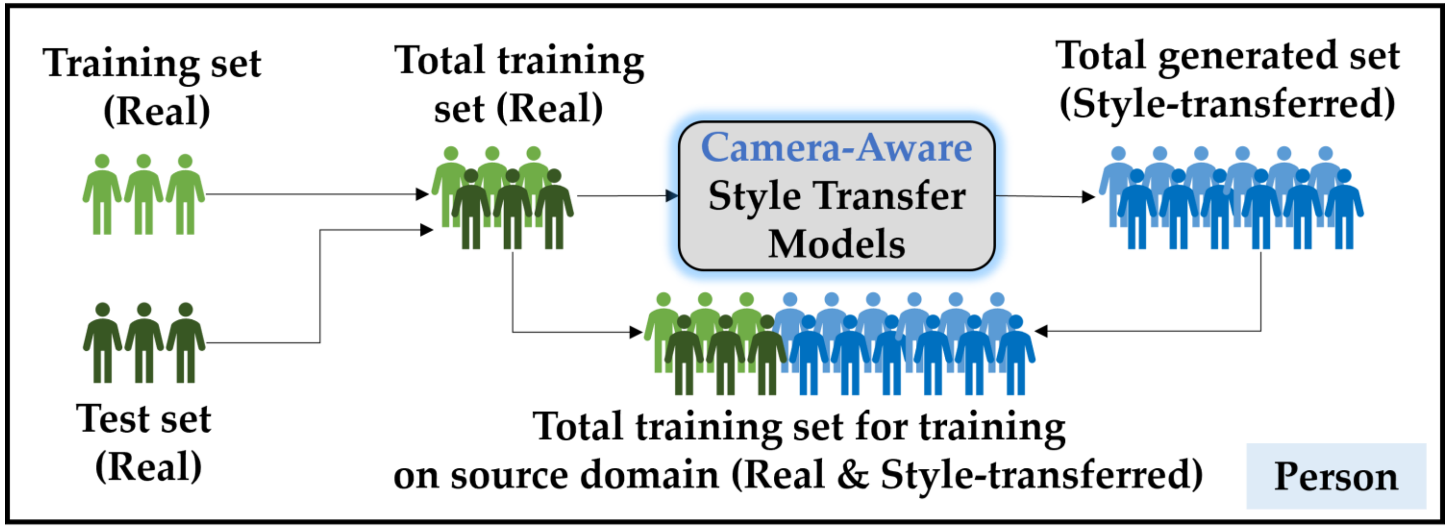
Proposed Methods:: CORE-ReID V2

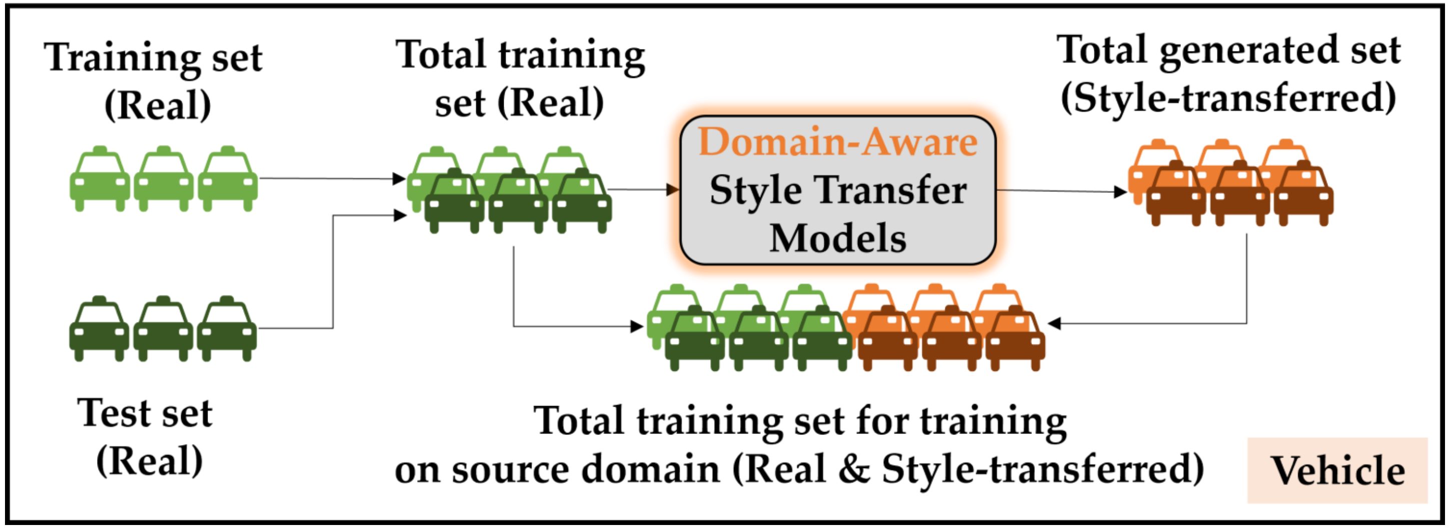
CORE-ReID V1's Limitations		CORE-ReID V2
Clustering limitations: random centroid initialization in K-means causes slow, unstable, and imbalanced results	▶	Enhanced clustering with greedy K-means++: uses greedy K-means++ for faster, more stable results
Feature fusion issue: the ECAB module enhanced only local features, ignore the global features	▶	Ensemble fusion++: enhances both local and global features
Restricted backbone support: only support heavy networks (ResNet 50, 101, 152)	▶	Flexible backbone support: supports lightweight backbones like ResNet18/34 for edge use

Methodology:: Pre-Training

- Pre-Training: Camera/Domain-aware style transfer on source dataset

Create the full training set for the source domain

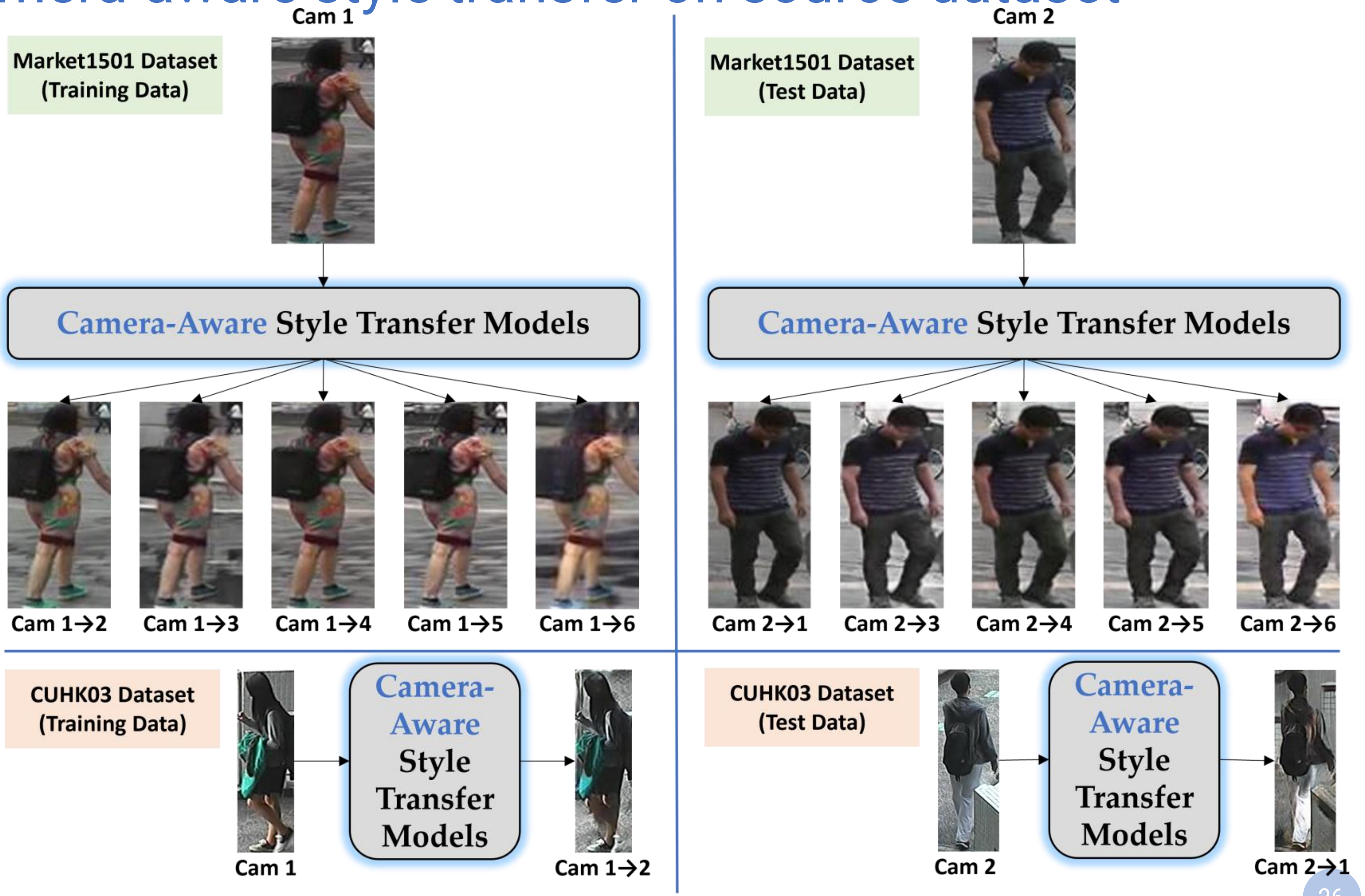




Methodology:: Pre-Training

Pre-Training: Camera-aware style transfer on source dataset

CycleGAN [13] was used to build Style Transfer Models

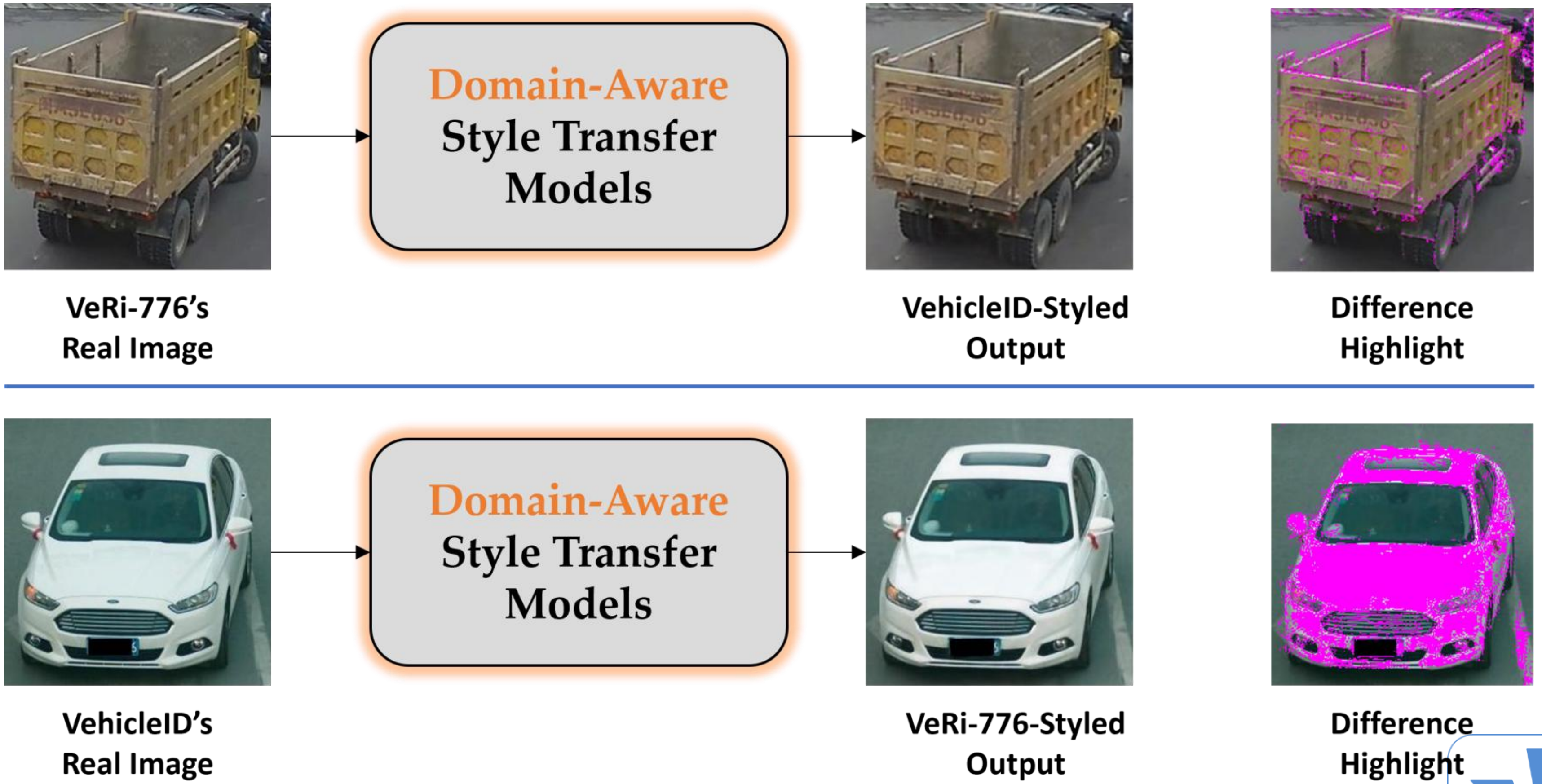


Iwate Prefectural University

Methodology:: Pre-Training

- Pre-Training: with some datasets not specifying the number of cameras, we applied Domain-aware style transfer approach.

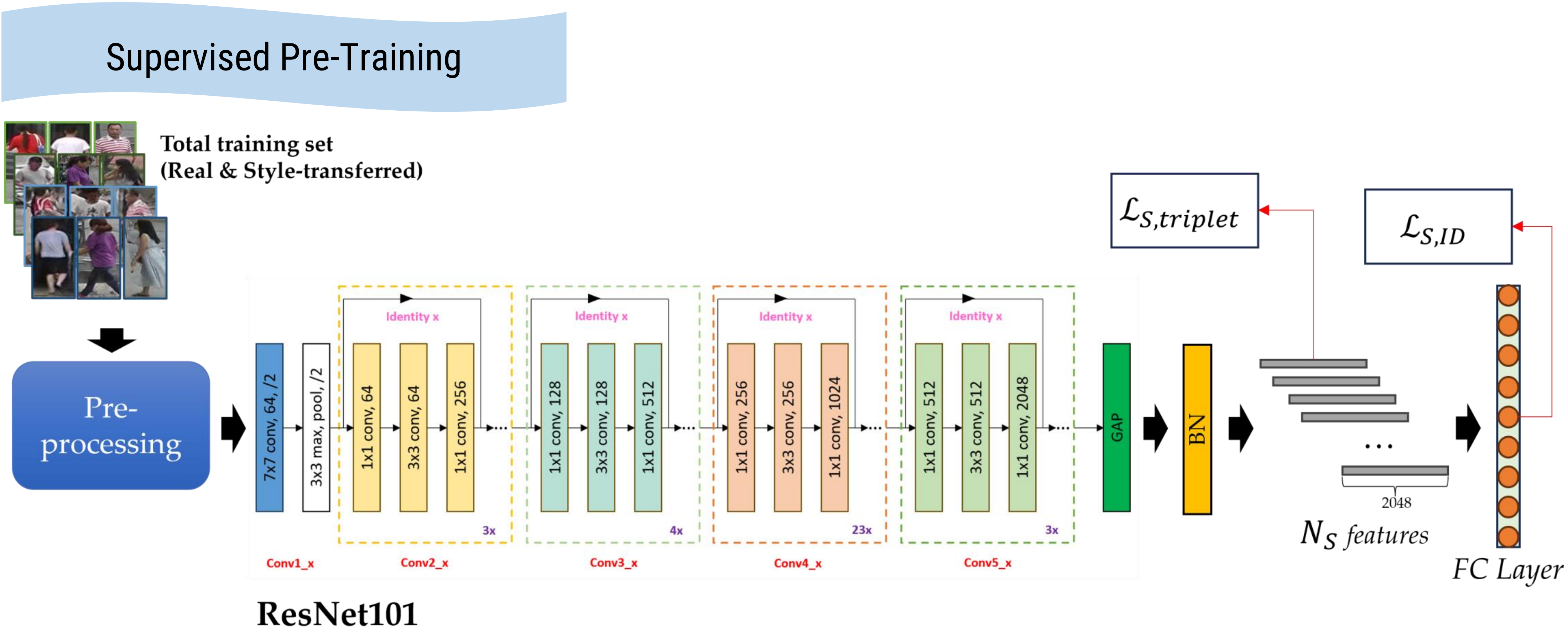
CycleGAN [13] was used to build Style Transfer Models



Iwate Prefectural University

Methodology:: Pre-Training

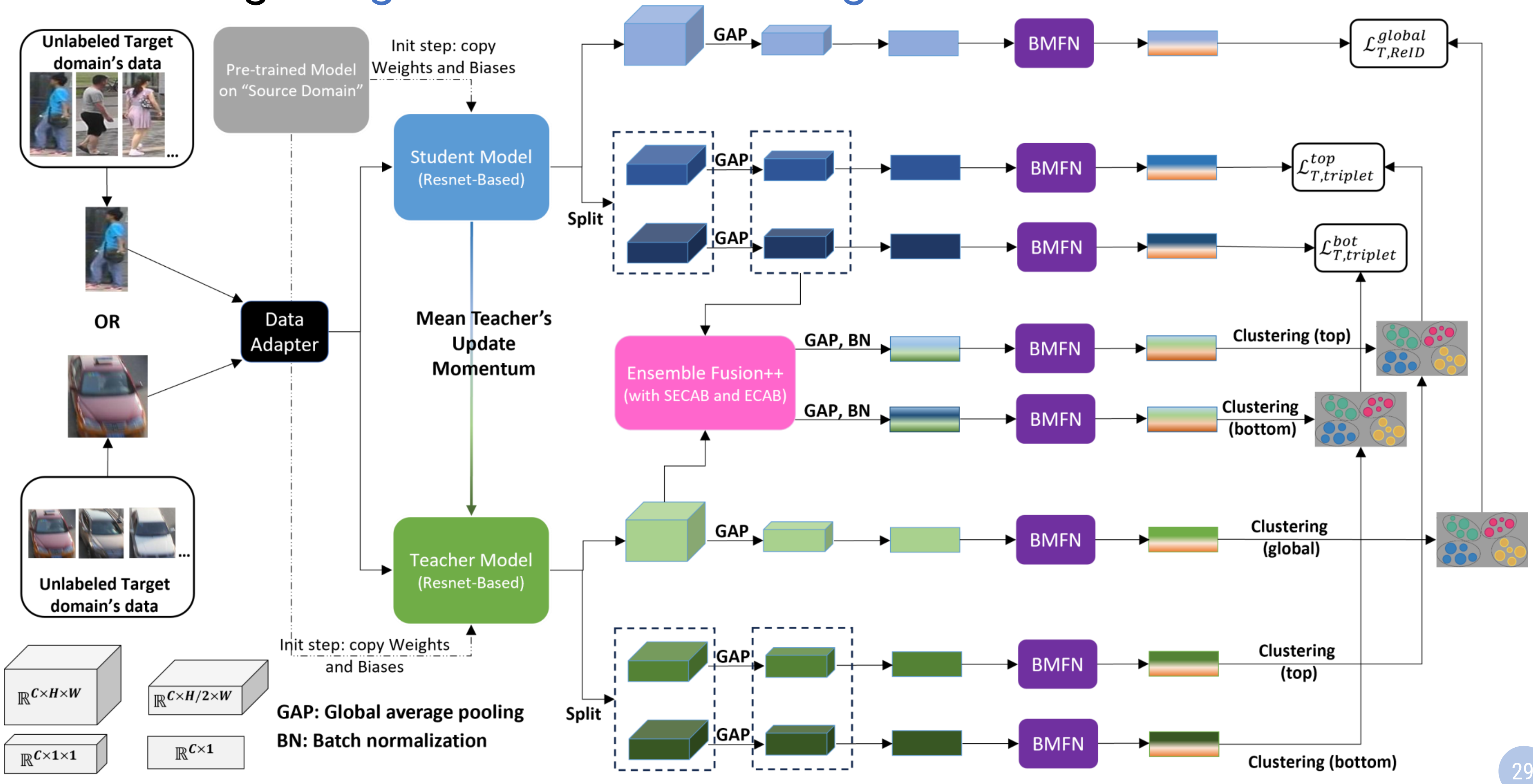
- Pre-Training: Source-domain pre-training



The overall training process in the fully supervised pre-training stage.
ResNet101 is used as the backbone in the training process.

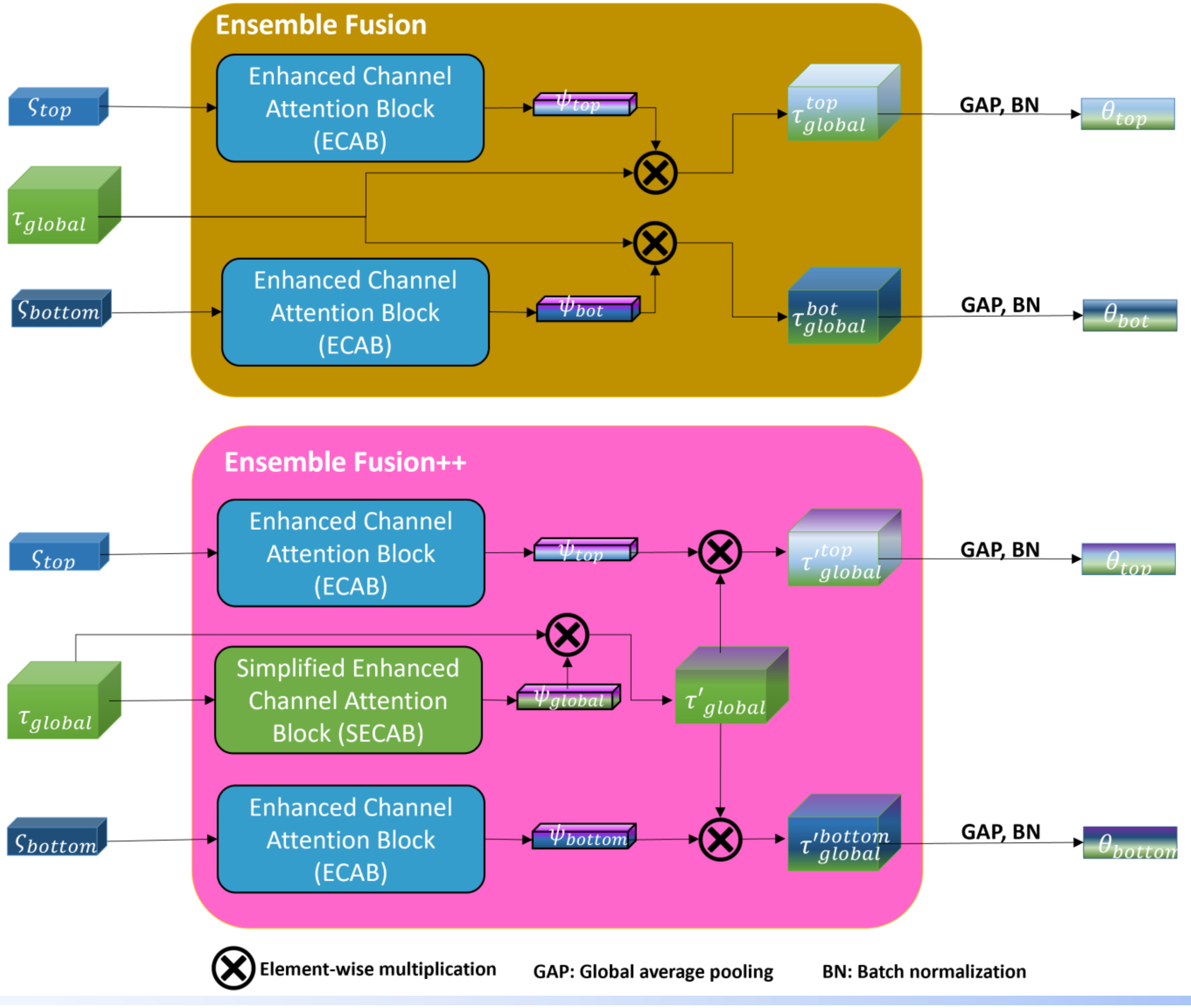
Methodology:: Fine Tuning

Fine Tuning: Target-domain fine-tuning



Fine Tuning: Target-domain fine-tuning

The comparison between Ensemble Fusion and proposed Ensemble Fusion++ component.



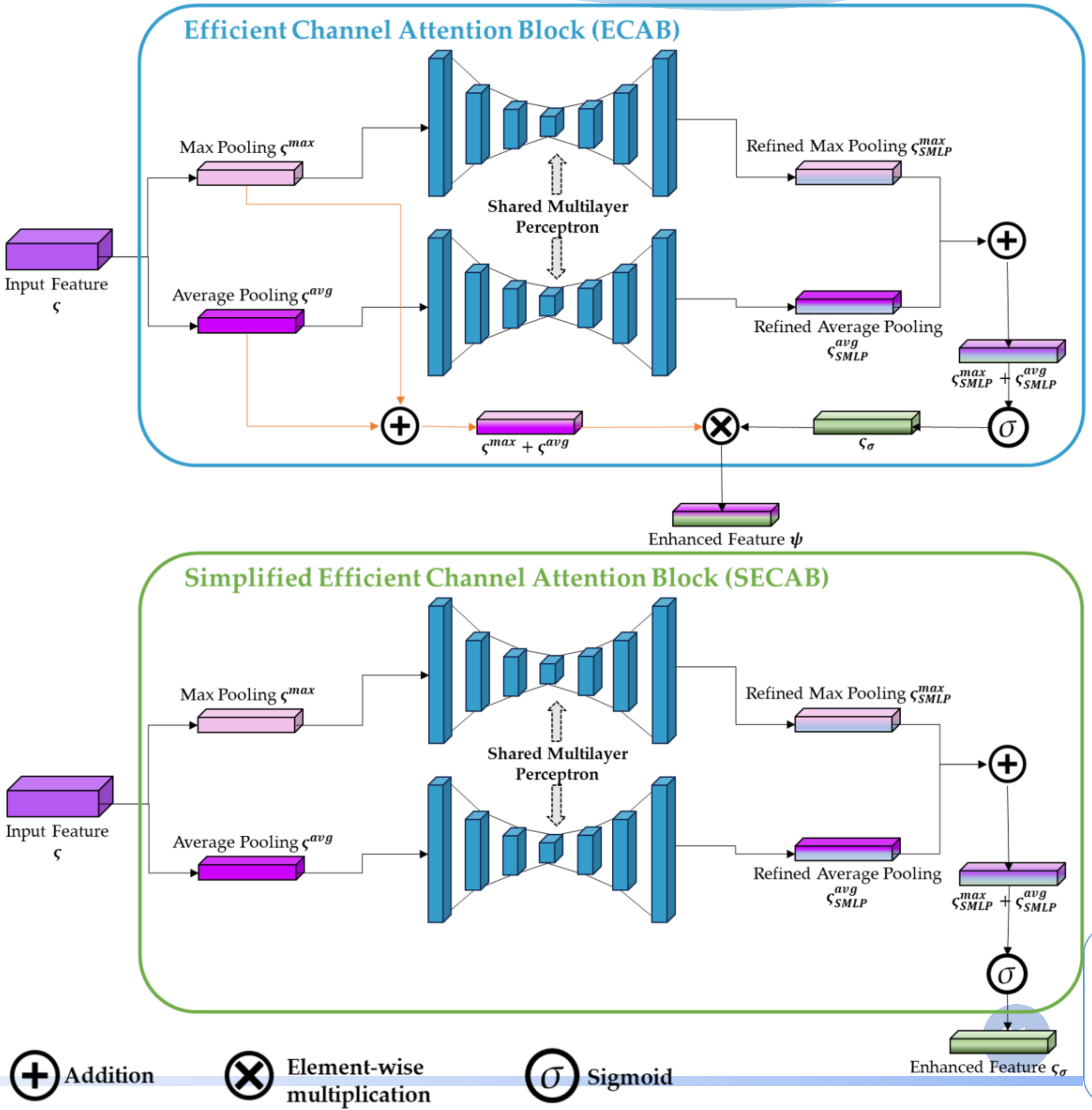
Iwate Prefectural University

Methodology:: Fine Tuning

Fine Tuning: Target-domain fine-tuning

Aspect	ECAB	SECAB
Target Use	Local feature vectors	Global feature map
Pooling	Adaptive Max + Avg Pooling	Adaptive Max + Avg Pooling
Attention Core	Shared Multilayer Perceptron	Same Shared Multilayer Perceptron
Output Processing	Attention map × (max + avg feature)	Attention map only
Residual Information Fusion (Later)	With refined global features	With original global features
Computational Cost	Higher (due to residual and additional element-wise operations)	Lower (no fusion step, lightweight on GPU)
Deployment Stage	Local-level features refinement	Global-level features refinement
Used in Ensemble Fusion	Yes	No
Used in Ensemble Fusion++	Yes	Yes

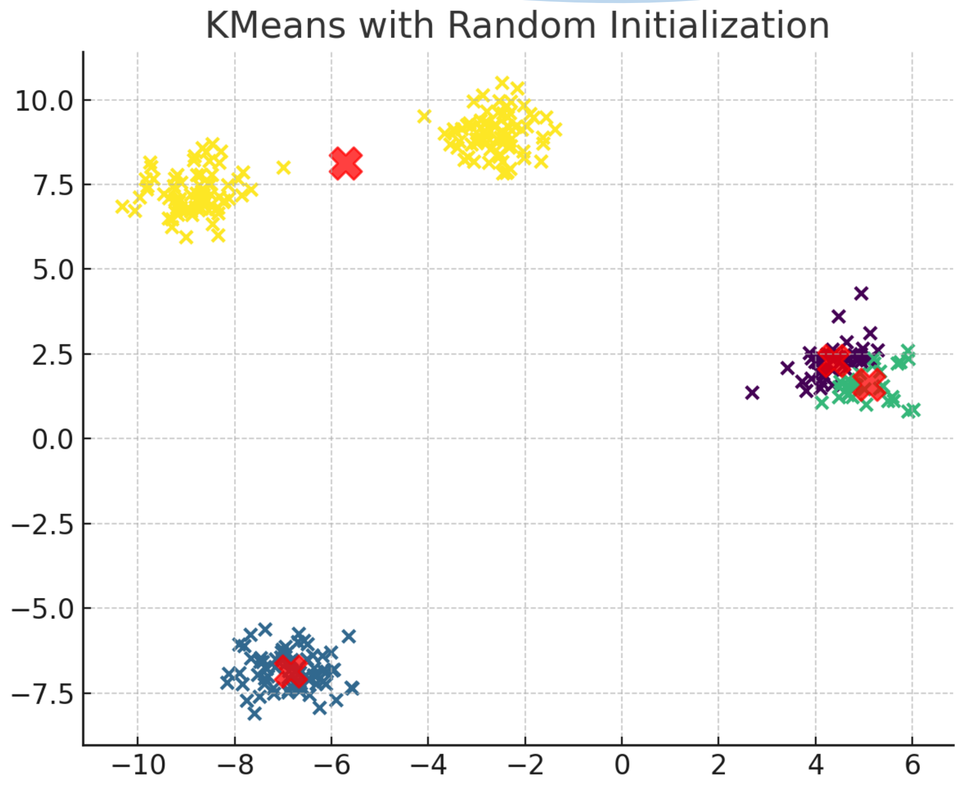
ECAB and SECAB designs



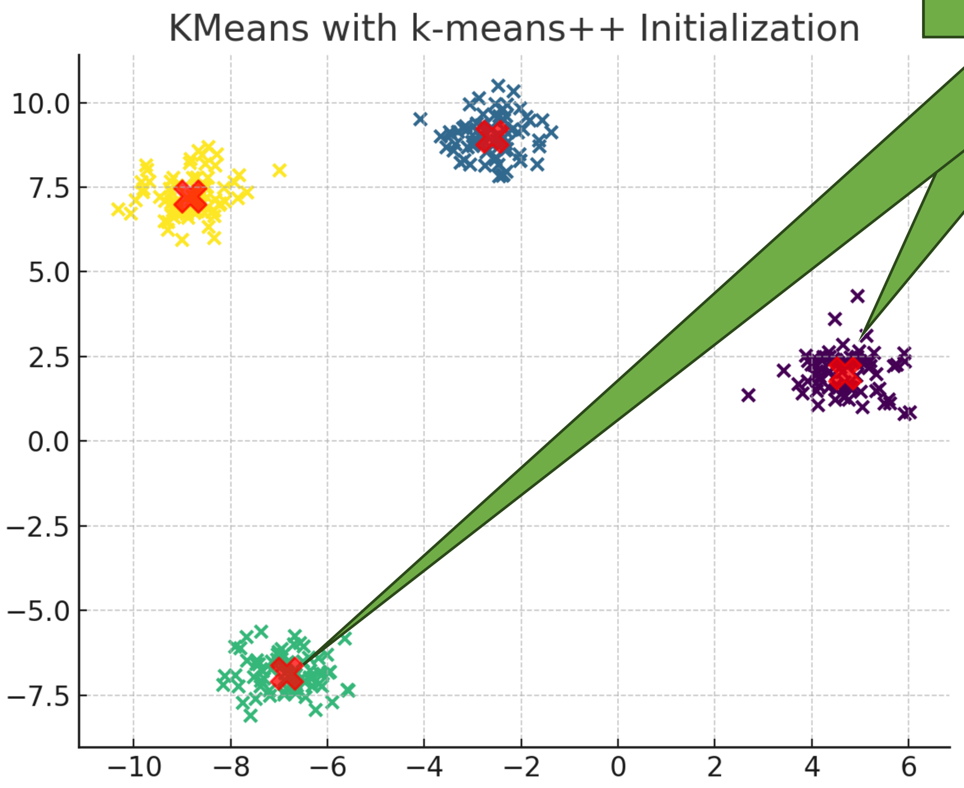
Methodology:: Fine Tuning

Fine Tuning: Target-domain fine-tuning

Greedy K-means++ initialization



Cluster centers are initialized randomly



The initialized clusters are much better

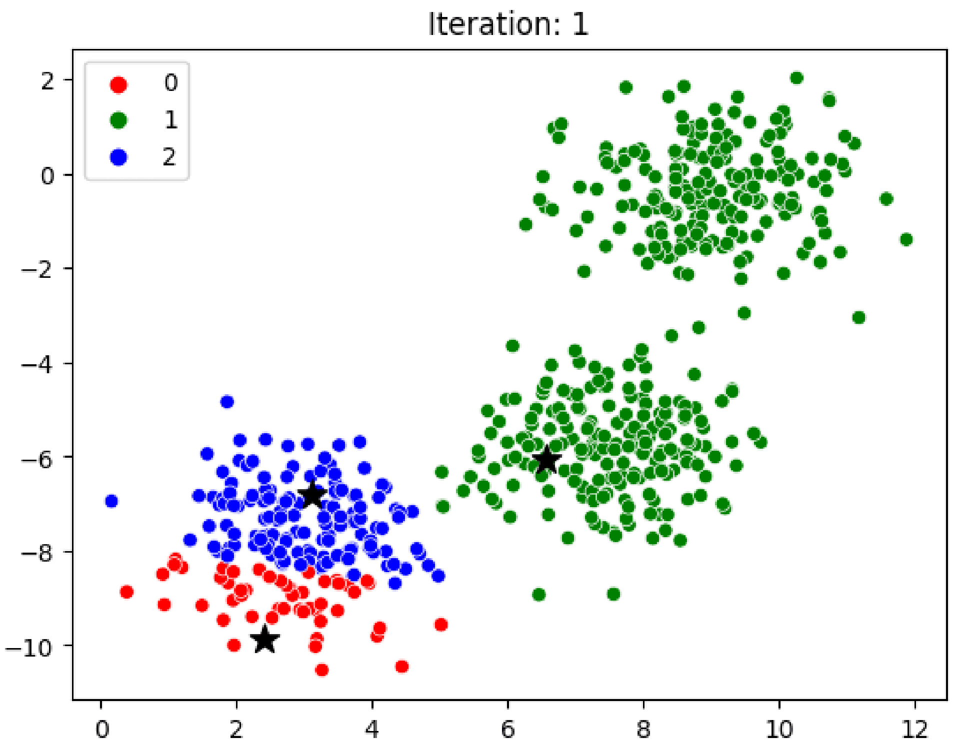
Cluster centers are initialized using a greedy strategy to maximize the initial separation.

Iwate Prefectural University

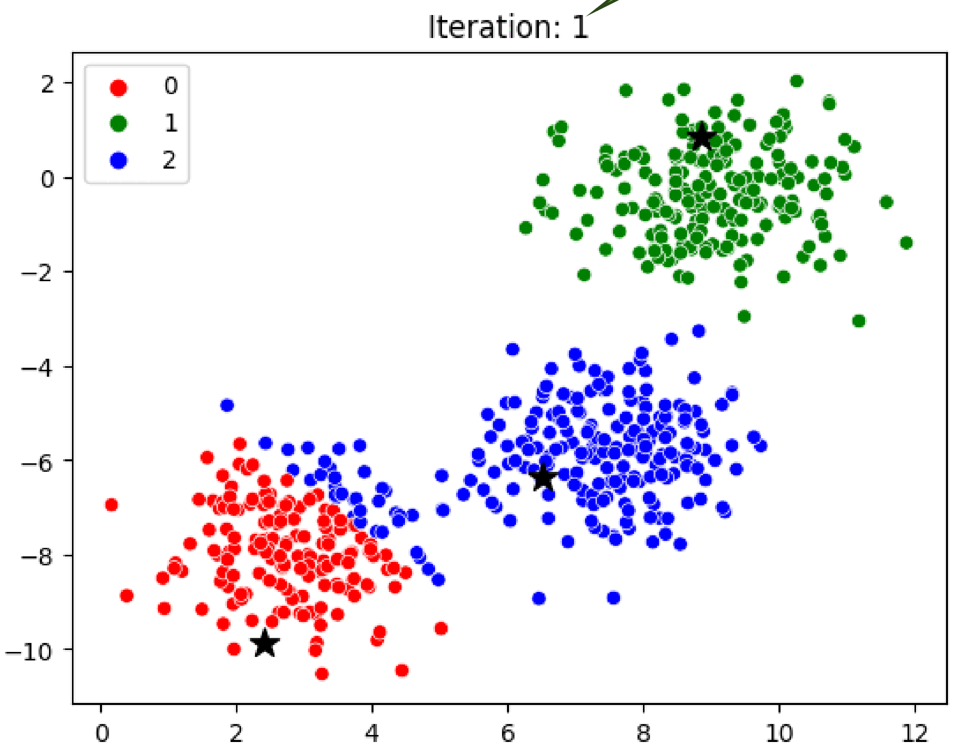
Fine Tuning: Target-domain fine-tuning

Greedy K-means++ initialization

The number of iterations is less (4 instead of 7)



Cluster centers are initialized randomly



Cluster centers are initialized using a greedy strategy to maximize the initial separation.

Results:: Evaluation Datasets

Six benchmark datasets

Category	Dataset	Location	Environment	Cameras	Training Set (ID/Image)	Test Set (ID/Image)	
						Gallery	Query
Person ReID	Market-1501	Tsinghua University, China	Outdoor campus	6	751/12,936	750/19,732	750/3,368
	CUHK03	Chinese University of Hong Kong	Indoor	2	767/7,365	700/5,332	700/1,400
	MSMT17	A university campus in China	Outdoor and indoor	15	1,401/32,621	3,060/82,161	3,060/11,659
Vehicle ReID	VeRi-776	A city in China	Surveillance cameras	20	576/37,778	200/11,579	200/1,678
	VehicleID	China (highways and streets)	Surveillance cameras	-	13,134/110,178	Test800: 800/800 Test1600: 1,600/1,600 Test2400: 2,400/2,400	Test800: 800/6,532 Test1600: 1,600/11,395 Test2400: 2,400/17,638
	VERI-Wild	Multiple cities in China	Surveillance cameras in a wide area	174	30,671/277,794	Test3000: 3,000/38,816 Test5000: 5,000/64,389 Test10000: 10,000/128,517	Test3000: 3,000/3,000 Test5000: 5,000/5,000 Test10000: 10,000/10,000

Results:: Evaluation Datasets

Six benchmark datasets

Market-1501 [16]

CUHK03 [17]

MSMT17 [18]

Veri-776 [19]

VehicleID [20]

VERI-Wild [21]

Results:: Market → CUHK & CUHK → Market

		Market → CUHK				CUHK → Market			
Method	Reference	mAP	R-1	R-5	R-10	mAP	R-1	R-5	R-10
SNR [a] [96]	CVPR 2020	17.5	17.1	-	-	52.4	77.8	-	-
UDAR [14]	PR 2020	20.9	20.3	-	-	56.6	77.1	-	-
$QAConv_{50}$ [a] [97]	ECCV 2020	32.9	33.3	-	-	66.5	85.0	-	-
M^3L [a] [98]	CVPR 2021	35.7	36.5	-	-	62.4	82.7	-	-
MetaBIN [a] [99]	CVPR 2021	43.0	43.1	-	-	67.2	84.5	-	-
DFH-Baseline [100]	CVPR 2022	10.2	11.2	-	-	13.2	31.1	-	-
DFH [a] [100]	CVPR 2022	27.2	30.5	-	-	31.3	56.5	-	-
META [a] [101]	ECCV 2022	47.1	46.2	-	-	76.5	90.5	-	-
ACL [a] [102]	ECCV 2022	49.4	50.1	-	-	76.8	90.6	-	-
RCFA [103]	Electronics 2023	17.7	18.5	33.6	43.4	34.5	63.3	78.8	83.9
CRS [104]	JSJTU 2023	-	-	-	-	65.3	82.5	93.0	95.9
MTI [105]	JVCIR 2024	16.3	16.2	-	-	-	-	-	-
PAOA+ [a] [106]	WACV 2024	50.3	50.9	-	-	77.9	91.4	-	-
Baseline (CORE-ReID) [11]	Software 2024	62.9	61.0	79.6	87.2	83.6	93.6	97.3	98.7
Direct Transfer	Ours	23.9	24.6	40.3	48.9	35.5	63.3	77.8	83.2
CORE-ReID V2 Tiny (ResNet18)	Ours	33.0	31.9	48.9	59.1	60.3	83.4	91.8	94.7
CORE-ReID V2	Ours	**66.4**	**66.9**	**83.4**	**88.9**	**84.5**	**93.9**	**97.6**	**98.7**

Bold denotes the best while Underline indicates the second-best results. [a] indicates the method uses multiple source datasets.

Iwate Prefectural University

Results:: Market → MSMT & CUHK → MSMT

Method	Reference	Market → MSMT mAP	Market → MSMT R-1	Market → MSMT R-5	Market → MSMT R-10	CUHK → MSMT mAP	CUHK → MSMT R-1	CUHK → MSMT R-5	CUHK → MSMT R-10
NRMT [107]	ECCV 2020	19.8	43.7	56.5	62.2	-	-	-	-
DG-Net++ [87]	ECCV 2020	22.1	48.4	-	-	-	-	-	-
MMT [15]	ICLR 2020	22.9	52.5	-	-	13.5 [b]	30.9 [b]	44.4 [b]	51.1 [b]
UDAR [14]	PR 2020	12.0	30.5	-	-	11.3	29.6	-	-
Dual-Refinement [108]	ArXiv 2020	25.1	53.3	66.1	71.5	-	-	-	-
SNR [a] [96]	CVPR 2020	-	-	-	-	7.7	22.0	-	-
$QAConv_{50}$ [a] [97]	ECCV 2020	-	-	-	-	17.6	46.6	-	-
M^3L [a] [98]	CVPR 2021	-	-	-	-	17.4	38.6	-	-
MetaBIN [a] [99]	CVPR 2021	-	-	-	-	18.8	41.2	-	-
RDSBN [109]	CVPR 2021	30.9	61.2	73.1	77.4	-	-	-	-
ClonedPerson [110]	CVPR 2022	14.6	41.0	-	-	13.4	42.3	-	-
META [a] [101]	ECCV 2022	-	-	-	-	24.4	52.1	-	-
ACL [a] [102]	ECCV 2022	-	-	-	-	21.7	47.3	-	-
CLM-Net [111]	NCA 2022	29.0	56.6	69.0	74.3	-	-	-	-
CRS [104]	JSJTU 2023	22.9	43.6	56.3	62.7	22.2	42.5	55.7	62.4
HDNet [112]	IJMLC 2023	25.9	53.4	66.4	72.1	-	-	-	-
DDNet [113]	AI 2023	28.5	59.3	72.1	76.8	-	-	-	-
CaCL [114]	ICCV 2023	36.5	66.6	75.3	80.1	-	-	-	-
PAOA+ [a] [106]	WACV 2024	-	-	-	-	26.0	52.8	-	-
OUDA [115]	WACV 2024	20.2	46.1	-	-	-	-	-	-
M-BDA [116]	VCIR 2024	26.7	51.4	64.3	68.7	-	-	-	-
UMDA [117]	VCIR 2024	32.7	62.4	72.7	78.4	-	-	-	-
Baseline (CORE-ReID) [11]	Software 2024	<u>41.9</u>	<u>69.5</u>	<u>80.3</u>	<u>84.4</u>	<u>40.4</u>	<u>67.3</u>	<u>79.0</u>	<u>83.1</u>
Direct Transfer	Ours	11.7	30.2	42.9	48.0	35.5	63.3	77.8	82.7
CORE-ReID V2 Tiny (ResNet18)	Ours	35.8	64.7	76.6	80.8	18.8	44.2	57.1	62.3
CORE-ReID V2	Ours	**44.1**	**71.3**	**82.4**	**86.0**	**40.7**	**68.7**	**79.7**	**83.4**

Bold denotes the best while <u>Underline</u> indicates the second-best results. [a] indicates the method uses multiple source datasets, [b] denotes the implementation is based on the author's code.

Results:: VehicleID → VeRi-776

Method	Reference	VehicleID → VeRi-776			
		mAP	R-1	R-5	R-10
FACT [1]	ECCV 2016	18.75	52.21	72.88	-
PUL [42]	ACM 2018	17.06	55.24	67.34	-
SPGAN [66]	CVPR 2018	16.4	57.4	70.0	75.6
VR-PROUD [118]	PR 2019	22.75	55.78	70.02	-
ECN [119]	CVPR 2019	20.06	57.41	70.53	-
MMT [15]	ICLR 2020	35.3	74.6	82.6	-
SPCL [44]	NIPS 2020	38.9	80.4	86.8	-
PAL [120]	IJCAI 2020	42.04	68.17	79.91	-
UDAR [14]	PR 2020	35.80	76.90	85.80	<u>89.00</u>
ML [121]	ICME 2021	36.90	77.80	85.50	-
PLM [122]	Sci.China 2022	47.37	77.59	87.00	-
VDAF [123]	MTA 2023	24.86	46.32	55.17	-
CSP+FCD [124]	Elec 2023	45.60	74.30	83.70	-
MGR-GCL [5]	ArXiv 2024	48.73	<u>79.29</u>	87.95	-
MATNet+DMDU [125]	ArXiv 2024	<u>49.25</u>	79.13	<u>88.97</u>	-
Baseline	Ours	47.70	78.12	86.23	88.14
Direct Transfer	Ours	22.71	62.04	71.79	76.32
CORE-ReID V2 Tiny (ResNet18)	Ours	40.17	73.00	81.41	85.40
CORE-ReID V2	Ours	**49.50**	**80.15**	**89.05**	**90.29**

Bold values represent the best results while <u>Underline values</u> indicate the second-best performance. Baseline is *CORE-ReID V1* method.

Iwate Prefectural University

Results:: VehicleID → VERI-Wild

		VehicleID → VERI-Wild											
		Test3000				Test5000				Test10000			
Method	**Reference**	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
SPGAN [66]	CVPR 2018	24.1	59.1	76.2	-	21.6	55.0	74.5	-	17.5	47.4	66.1	-
ECN [119]	CVPR 2019	34.7	73.4	88.8	-	30.6	68.6	84.6	-	24.7	61.0	78.2	-
MMT [15]	ICLR 2020	27.7	55.6	77.4	-	23.6	47.7	71.5	-	18.0	40.2	65.0	-
SPCL [44]	NIPS 2020	25.1	48.8	72.8	-	21.5	42.0	66.1	-	16.6	32.7	55.7	-
UDAR [14]	PR 2020	30.0	68.4	85.3	-	26.2	62.5	81.8	-	20.8	53.7	73.9	-
AE [126]	CCA 2020	29.9	87.0	68.5	-	26.2	61.8	81.5	-	20.9	53.1	73.7	-
DLVL [18]	Elec 2024	31.4	59.9	80.7	-	27.3	51.9	74.9	-	21.7	41.8	65.8	-
Baseline	Ours	<u>39.8</u>	<u>75.2</u>	<u>89.3</u>	<u>91.6</u>	<u>34.5</u>	69.6	<u>81.7</u>	<u>88.7</u>	<u>26.8</u>	<u>61.1</u>	<u>79.6</u>	<u>81.3</u>
Direct Transfer	Ours	20.9	48.2	64.3	70.7	18.9	44.3	60.9	66.9	15.6	38.0	53.3	59.8
CORE-ReID V2 Tiny (ResNet18)	Ours	28.6	56.5	74.9	80.2	23.1	52.1	70.6	78.4	19.9	48.1	66.3	74.6
CORE-ReID V2	Ours	**40.2**	**76.6**	**90.2**	**92.1**	**34.9**	**70.2**	**86.2**	**89.3**	**27.8**	**62.1**	**79.8**	**82.3**

Bold values represent the best results while <u>Underline values</u> indicate the second-best performance. Baseline is *CORE-ReID V1* method.

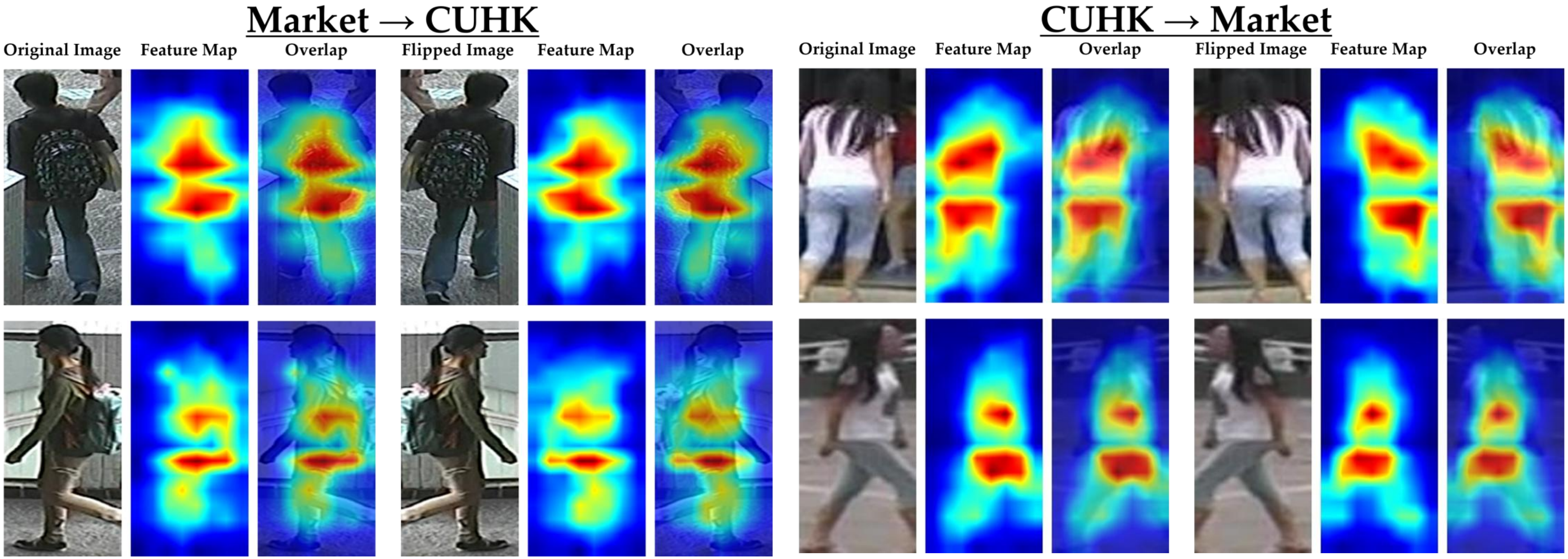
Results:: VeRi-776 → VehicleID

Method	Reference	VeRi-776 → VehicleID Test800				VeRi-776 → VehicleID Test1600				VeRi-776 → VehicleID Test2400			
		mAP	R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP	R-1	R-5	R-10
FACT [1]	ECCV 2016	-	49.53	67.96	-	-	44.63	64.19	-	-	39.91	60.49	-
Mixed Diff+CCL [84]	CVPR 2016	-	49.00	73.50	-	-	42.80	66.80	-	-	38.20	61.60	-
PUL [42]	ACM 2018	43.90	40.03	56.03	-	37.68	33.83	49.72	-	34.71	30.90	47.18	-
PAL [120]	IJCAI 2020	53.50	50.25	64.91	-	48.05	44.25	60.95	-	45.14	41.08	59.12	-
UDAR [14]	PR 2020	59.60	54.00	66.10	72.01	55.30	48.10	64.10	70.20	52.90	45.20	62.60	69.14
ML [121]	ICME 2021	61.60	54.80	69.20	-	48.70	40.30	57.70	-	45.00	36.50	54.10	-
PLM [122]	Sci.China 2022	54.85	51.23	67.11	-	49.41	45.40	63.37	-	46.00	41.73	60.94	-
CSP+FCD [124]	Elec 2023	51.90	54.40	67.40	-	46.50	52.70	65.60	-	42.70	45.90	60.30	-
VDAF [123]	MTA 2023	-	-	-	-	-	47.03	64.86	-	-	43.69	61.76	-
MGR-GCL [5]	ArXiv 2024	55.24	52.38	75.29	-	50.56	45.88	67.65	-	47.59	42.83	64.36	-
DMDU [125]	TITS 2024	61.83	55.61	68.25	-	56.73	53.28	63.56	-	53.97	47.59	61.85	-
Baseline	Ours	64.28	56.16	74.55	81.15	60.02	51.84	71.62	78.08	56.15	47.85	66.89	75.27
Direct Transfer	Ours	61.28	53.50	69.81	76.13	57.23	48.57	67.05	73.77	52.31	44.04	61.08	68.60
CORE-ReID V2 Tiny (ResNet18)	Ours	63.87	55.18	73.43	81.11	59.69	50.05	70.88	77.75	55.14	45.99	65.07	73.54
CORE-ReID V2	Ours	**67.04**	**58.32**	**76.51**	**84.32**	**63.02**	**53.49**	**74.36**	**81.85**	**57.99**	**48.62**	**68.30**	**77.11**

Bold values represent the best results while Underline values indicate the second-best performance. Baseline is *CORE-ReID V1* method.

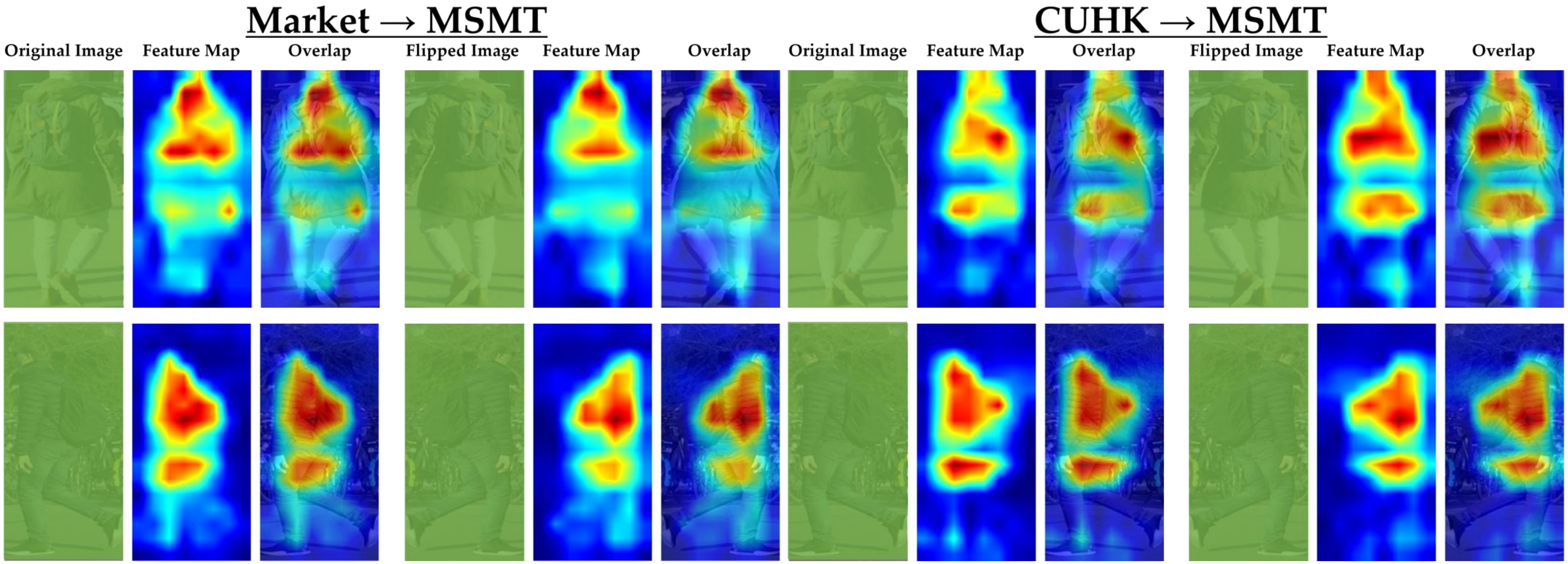
Ablation Study:: Feature Maps Visualization

The feature map, visualized using Grad-CAM [22] at the global feature level, highlights important features of each individual, represented as heatmaps.



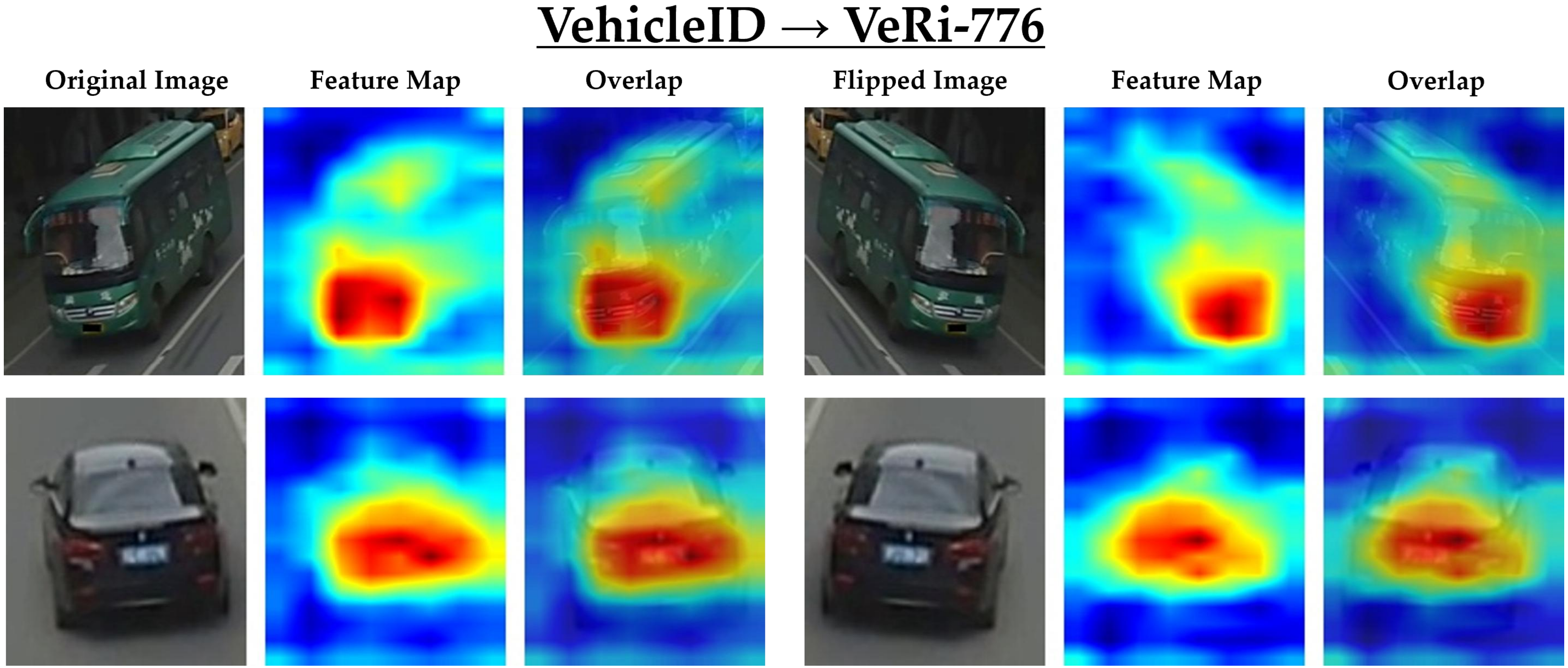
Ablation Study:: Feature Maps Visualization

The feature map, visualized using Grad-CAM at the global feature level, highlights important features of each individual, represented as heatmaps.



Ablation Study:: Feature Maps Visualization

The feature map, visualized using Grad-CAM at the global feature level, highlights important features of each individual, represented as heatmaps.

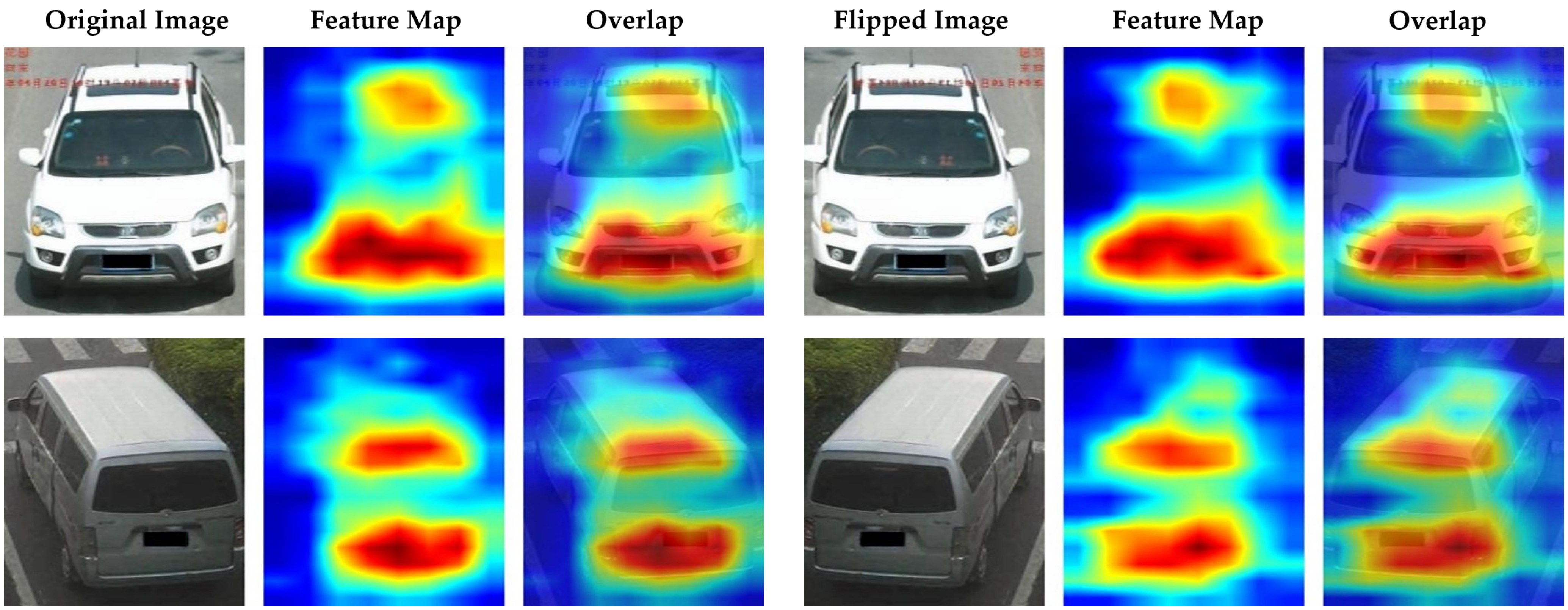


Iwate Prefectural University

Ablation Study:: Feature Maps Visualization

The feature map, visualized using Grad-CAM at the global feature level, highlights important features of each individual, represented as heatmaps.

VeRi-776 → VehicleID



Ablation Study:: K-means Clustering Settings

The K-means algorithm was employed for clustering to generate pseudo-labels in the target domain.

Person ReID	Market → CUHK				CUHK → Market			
Number of Clusters	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 500$)	44.4	43.2	65.3	76.4	69.4	86.8	94.9	96.7
Ours ($M_{T,j} = 700$)	57.8	59.1	76.1	83.6	81.7	92.7	97.1	98.1
Ours ($M_{T,j} = 900$)	**66.4**	**66.9**	**83.4**	**88.9**	**84.5**	**93.9**	**97.6**	**98.7**
Person ReID	**Market → MSMT**				**CUHK → MSMT**			
Number of Clusters	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 2000$)	**44.1**	**71.3**	**82.4**	**86.0**	**40.68**	**68.66**	**79.74**	**83.36**
Ours ($M_{T,j} = 2500$)	41.1	68.9	80.5	84.2	38.91	67.26	78.97	82.80
Ours ($M_{T,j} = 3000$)	38.9	67.2	79.0	83.2	35.8	64.7	76.6	80.8
Vechile ReID	**VehicleID → VeRi-776**				**VeRi-776 → VehicleID Small**			
Number of Clusters	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 500$)	49.50	**80.15**	**89.05**	**90.29**	66.60	58.20	75.90	83.70
Ours ($M_{T,j} = 700$)	**49.63**	79.14	86.65	89.69	**67.04**	**58.32**	**76.51**	**84.32**
Ours ($M_{T,j} = 900$)	48.61	79.02	86.29	89.15	66.70	57.50	77.60	84.20

Experimental results on different settings of number of pseudo identities in K-means clustering algorithm.
Bold denotes the best results.

Ablation Study:: Greedy K-means++ Initialization

The experimental results comparing greedy K-means++ initialization with a random approach.

Person ReID	Market → CUHK				CUHK → Market			
Number of Clusters	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 500$)	44.4	43.2	65.3	76.4	69.4	86.8	94.9	96.7
Ours ($M_{T,j} = 700$)	57.8	59.1	76.1	83.6	81.7	92.7	97.1	98.1
Ours ($M_{T,j} = 900$)	**66.4**	**66.9**	**83.4**	**88.9**	**84.5**	**93.9**	**97.6**	**98.7**
Person ReID	**Market → MSMT**				**CUHK → MSMT**			
Number of Clusters	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 2000$)	**44.1**	**71.3**	**82.4**	**86.0**	**40.68**	**68.66**	**79.74**	**83.36**
Ours ($M_{T,j} = 2500$)	41.1	68.9	80.5	84.2	38.91	67.26	78.97	82.80
Ours ($M_{T,j} = 3000$)	38.9	67.2	79.0	83.2	35.8	64.7	76.6	80.8
Vechile ReID	**VehicleID → VeRi-776**				**VeRi-776 → VehicleID Small**			
Number of Clusters	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 500$)	49.50	**80.15**	**89.05**	**90.29**	66.60	58.20	75.90	83.70
Ours ($M_{T,j} = 700$)	**49.63**	79.14	86.65	89.69	**67.04**	**58.32**	**76.51**	**84.32**
Ours ($M_{T,j} = 900$)	48.61	79.02	86.29	89.15	66.70	57.50	77.60	84.20

Experimental results on different settings of number of pseudo identities in K-means clustering algorithm.
Bold denotes the best results.

Ablation Study:: SECAB Configuration

To validate the effectiveness of SECAB, we conduct an experiment by removing it from our network.

Person ReID	Market → CUHK ($M_{T,j}$ = 900)				CUHK → Market ($M_{T,j}$ = 900)			
Method	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours (without SECAB)	65.0	65.1	82.6	87.6	83.9	93.7	97.4	98.6
Ours (with SECAB)	**66.4**	**66.9**	**83.4**	**88.9**	**84.5**	**93.9**	**97.6**	**98.7**
Person ReID	**Market → MSMT ($M_{T,j}$ = 2000)**				**CUHK → MSMT ($M_{T,j}$ = 2000)**			
Method	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours (without SECAB)	43.2	70.3	81.8	85.2	40.5	68.0	79.2	83.1
Ours (with SECAB)	**44.1**	**71.3**	**82.4**	**86.0**	**40.7**	**68.7**	**79.7**	**83.4**
Vehicle ReID	**VehicleID → VeRi-776 ($M_{T,j}$ = 500)**				**VeRi-776 → VehicleID Small ($M_{T,j}$ = 700)**			
Method	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours (without SECAB)	48.03	78.92	87.61	88.93	65.14	57.02	75.56	82.97
Ours (with SECAB)	**49.50**	**80.15**	**89.05**	**90.29**	**67.04**	**58.32**	**76.51**	**84.32**

Experimental results on different settings of number of pseudo identities in K-means clustering algorithm.
Bold denotes the best results.

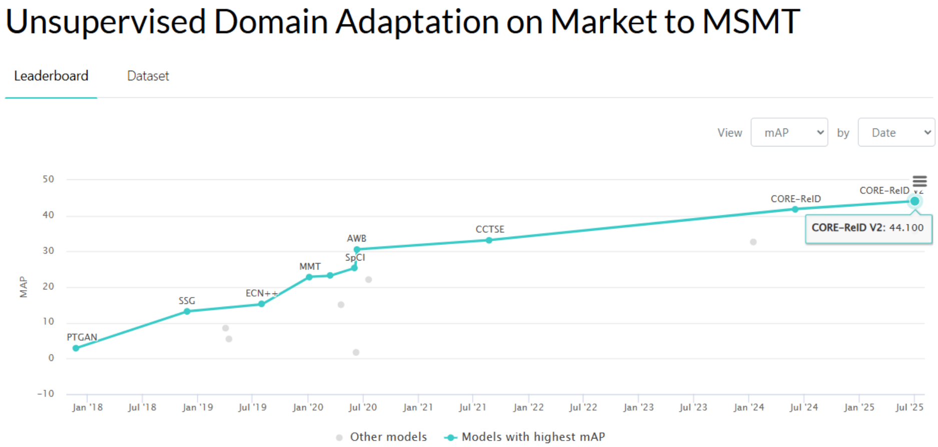
Ablation Study:: Computational Complexity Analysis

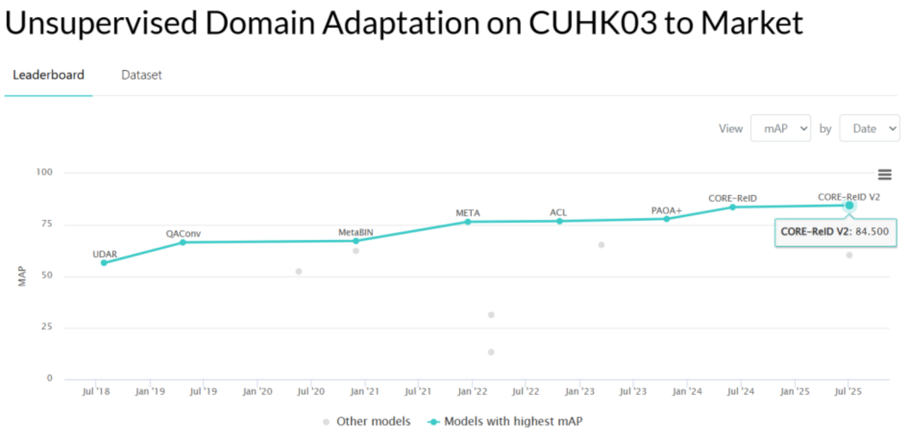
We include a comparative table of commonly used Res-Net backbones in terms of parameters and Giga Floating-Point Operations per Second (GFLOPs).

CORE-ReID V2 with Backbone	Parameters (millions)	GFLOPs (per image)	Image Size	FPS (using 1 Quadro RTX 8000 GPU)
ResNet-18	12.97 M	1.18	128x256	254
ResNet-34	23.08 M	2.35	128x256	185
ResNet-50	46.62 M	5.10	128x256	144
ResNet-101	65.61 M	7.58	128x256	87
ResNet-152	81.26 M	10.61	128x256	61

The clustering parameter values ($M_{T,j}$) is carried out from the study of K-means clustering settings.
Bold denotes the best results.

Ablation Study:: Benchmark on PaperWithCode





#	Source dataset	Target dataset	Paper with code (CORE-ReID)	Rank
1	CUHK03	Market-1501	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-cuhk03-to-1	Top1
2	CUHK03	MSMT17	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-cuhk03-to	Top1
3	Market-1501	CUHK03	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-market-to-6	Top1
4	Market-1501	MSMT17	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-market-to-1	Top1
5	VehicleID	Veri-776	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-vehicleid	Top1
6	VehicleID	VERI-Wild	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-vehicleid-1 https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-vehicleid-2 https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-vehicleid-3	Top1
7	Veri-776	VehicleID	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-veri-776-to https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-veri-776-to-1 https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-veri-776-to-2	Top1

Summary

CORE-ReID V1 & V2

Category	CORE-ReID V1		CORE-ReID V2
	Current Status	Drawbacks/ Issues	
Applied Domain	Person ReID	Only support Person ReID.	Expansion from Person ReID to Vehicle ReID and further Object ReID.
Synthetic Data Generation	Camera-Aware Style Transfer	Do not work in case the number of cameras is not specified.	Camera-Aware Style Transfer and Domain-Aware Style Transfer (for the case the number of cameras is not specified).
Data Augmentation	Random gray scale patch replacement	Only replace random gray scale patch in the image locally.	Locally gray scale patch replacement and global gray scale conversion.
K-Means Clustering	Random initialization	Problems from random initialization (1) Poor centroid placement (2) Slow convergence (3) Stuck in local minima (4) High variance in results (5) Imbalanced cluster sizes	Greedy K-Means++ initialization helps: (1) Selects centroids with optimized spread (2) Minimizes redundancy, requiring fewer iterations (3) Improves initialization stability (4) Reduces randomness and provides consistent clusters (5) Ensures better centroid distribution
Ensemble Fusion	Ensemble Fusion with ECAB	Only the local features are enhanced in the Ensemble Fusion.	Ensemble Fusion++ (with ECAB and SECAB) helps enhance both local and global features.
Supported Backbones	ResNet50, 101, 152	Do not support small backbones such as ResNet18, 34.	ResNet18, 34, 50, 101, 152

Conclusion

Achievements and contributions

Advanced Data Augmentation Techniques: The framework integrates novel data augmentation strategies, such as Local Grayscale Patch Replacement and Random Image-to-Grayscale Conversion for UDA task. These methods introduce diversity in the training data, enhancing the model's stability.

Dynamic and Flexible Backbone Support: CORE-ReID V2 extends compatibility to smaller backbone architectures, including ResNet18 and ResNet34, without compromising performance. This flexibility allows for deployment in re-source-constrained environments while maintaining high accuracy.

Expansion to Vehicle and further Object ReID: Unlike its predecessor, which focused solely on person re-identification, CORE-ReID V2 extends its scope to Ve-hicle Re-identification and further general Object Re-identification. This expansion demonstrates its versatility and adaptability across various domains.

Introduction of Ensemble Fusion++: The framework incorporates the SECAB into the global feature extraction pipeline to enhance feature representation by dynamically emphasizing informative channels, thereby improving discrimination between instances.

Achieved state-of-the-art (SOTA) performance and reduced the gap between supervised and unsupervised Person Re-ID.

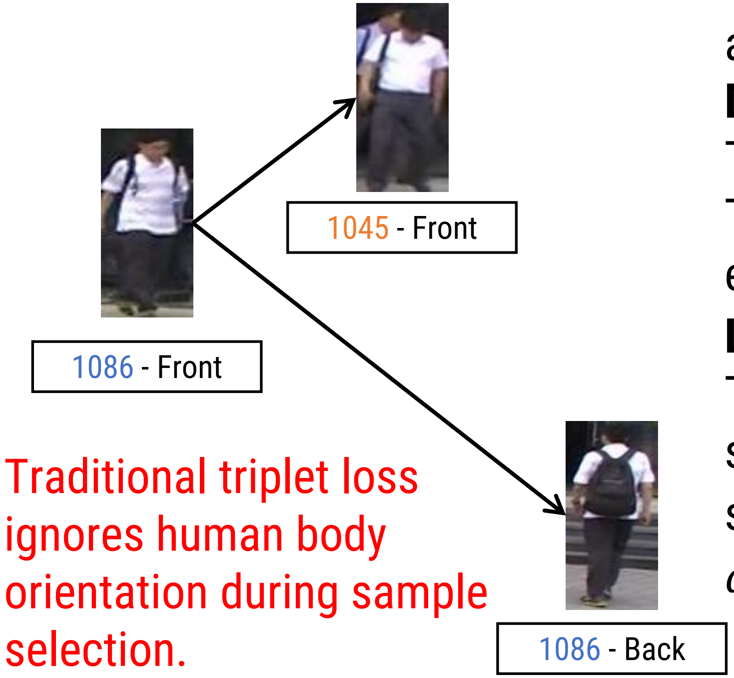
Future Work:: CORE-ReID V2

Limitations and Solutions

1. The scalability of the framework to other datasets remains unexplored
   - Explore other datasets (BV-Person, ENTIRe-ID, VRID-1, VRAI, Vehicle-Rear and V2I-CARLA,…)
2. Focus on Person and Vehicle ReID tasks also limits its exploration of broader Object ReID applications
   - Evaluate CORE-ReID V2 on general object, such as animal, product, or scene-specific ReID
3. Reliance on the quality of pseudo-labels makes it vulnerable to performance degradation in noisy or highly complex scenarios.
   - Explore advanced techniques, such as adversarial regularization, to mitigate the impact of noisy pseudo-labels.

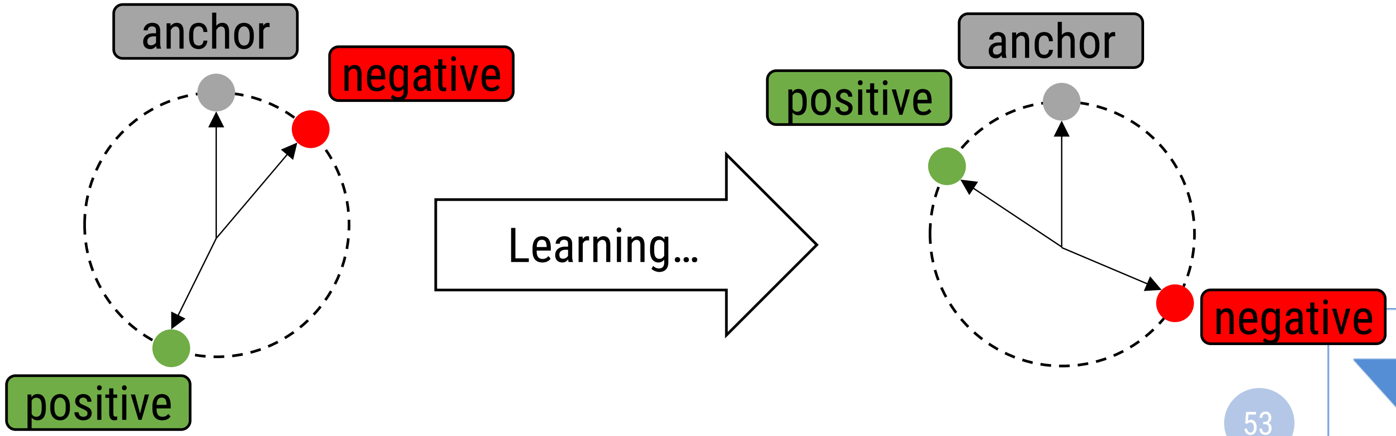
Future Work:: Orientation-aware triplet loss

Triplet Loss



Traditional triplet loss ignores human body orientation during sample selection.

Triplet Loss:
Triplet loss uses triplets of images: an **anchor** image, a **positive** image (same identity as anchor), and a **negative** image (different identity than anchor).
Distance Calculation:
The network calculates the feature embeddings for each image using a deep neural network. Then, it computes the Euclidean distance between the anchor's embedding and the positive embedding ($d_{positive}$) and the anchor's embedding and the negative embedding ($d_{negative}$).
Loss Function:
The core of triplet loss is to minimize the distance between the anchor and the positive sample ($d_{positive}$) while maximizing the distance between the anchor and the negative sample ($d_{negative}$).. This is often achieved with a margin (m) where $d_{positive} + m < d_{negative}$.



Future Work:: Orientation-aware triplet loss

Limitations and Solutions

Needs / Issues	Motivations	Solution
1. Traditional triplet loss ignores human body orientation during positive and negative sample selection. 2. Negative samples with vastly different orientations from the anchor (e.g., front vs. back view) are too easy, leading to ineffective training. 3. Models may be overfit to pose-related features rather than identity-discriminative features, reducing generalization in cross-camera or domain adaptation scenarios.)	1. To encourage the model to focus on identity cues, not just pose or orientation. 2. To improve the effectiveness of hard negative mining by making it orientation-aware. 3. To enhance robustness and generalization, especially in cross-camera or domain-adaptive person re-identification.	1. Modify the triplet sampling strategy: • Anchor: person image with a given orientation. • Positive: same person, not very similar orientation to the anchor. • Negative: different person, but with similar orientation to the anchor. 2. This forces the model to learn fine-grained, identity-specific features by removing orientation as a shortcut. 3. Body orientation is estimated using pose estimation models, classifiers, or metadata.

References

[1] Shlens, J. A Tutorial on Principal Component Analysis. arXiv 2014. DOI:10.48550/arXiv.1404.1100.

[2] Qu, L.; Pei, Y. A Comprehensive Review on Discriminant Analysis for Addressing Challenges of Class-Level Limitations, Small Sample Size, and Robustness. Processes 2024, 12, 1382. DOI:10.3390/pr12071382.

[3] Bromley, J.; Bentz, J. W.; Bottou, L.; Guyon, I. M.; LeCun, Y.; Moore, C.; Sackinger, E.; Shah, R. Signature Verification Using A "Siamese" Time Delay Neural Network. Int. J. Pattern Recognit. Artif. Intell. 1993, 7, 669–688. DOI:10.1142/S0218001493000339.

[4] Meng, Q.; Shichao, Z.; Zhida, H.; Feng, Z. MagFace: A Universal Representation for Face Recognition and Quality Assessment. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20-25 June 2021; pp. 14220–14229.

[5] Schroff, F; Kalenichenko, D.; Philbin, J. FaceNet: A Unified Embedding for Face Recognition and Clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 07-12 June 2015; pp. 815-823. DOI:10.1109/cvpr.2015.7298682.

[6] Mohan, D. D.; Jawade, B.; Setlur, S.; Govindaraj, V. Deep Metric Learning for Computer Vision: A Brief Overview. arXiv 2023. DOI: 10.48550/arXiv.2312.10046.

[7] Chen, T.; Kornblith, S; Norouzi, M; Hinton, G. A simple framework for contrastive learning of visual representations. In Proceedings of the 37th International Conference on Machine Learning, 13 July 2020; pp. 1597–1607.

[8] He, K.; Fan, H; Wu, Y.; Xie, S.; Girshick, R. Momentum Contrast for Unsupervised Visual Representation Learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13-19 June 2020; pp. 9726-9735. DOI:10.1109/CVPR42600.2020.00975.

[9] Chung, D.; Tahboub, K.; Delp, J. E. A Two Stream Siamese Convolutional Neural Network for Person Re-identification. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22-29 October 2017; pp. 1992-2000. DOI:10.1109/ICCV.2017.218.

[10] Mekhazni, D.; Dufau, M.; Desrosiers, C.; Pedersoli, M.; Granger, E. Camera Alignment and Weighted Contrastive Learning for Domain Adaptation in Video Person ReID. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 02-07 January 2023; pp. 1624-1633. DOI: 10.1109/WACV56688.2023.00167.

Iwate Prefectural University

References

[11] Ye, M.; Shen, J.; Lin, G.; Xiang, T.; Shao, L.; Hoi, S.C.H. Deep Learning for Person Re-Identification: A Survey and Outlook. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 2872–2893.

[12] Fu, Y.; Wei, Y.; Wang, G.; Zhou, Y; Shi, H.; Uiuc U. Self-Similarity Grouping: A Simple Unsupervised Cross Domain Adaptation Approach for Person Re-Identification. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 27 October 2019 - 02 November 2019; pp. 6111-6120. DOI:10.1109/ICCV.2019.00621.

[13] Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In 2Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.

[14] Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.

[15] Grunau, C.; Özüdoğru, A.A.; Rozhoň, V.; Tětek, J. A Nearly Tight Analysis of Greedy K-means++. In Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), Florence, Italy, 22–25 January 2023; pp. 1012–1070.

[16] Zheng, L.; Shen, L.; Tian, L.; Wang, S.; Wang, J.; Tian, Q. Scalable Person Re-identification: A Benchmark. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1116–1124.

[17] Li, W.; Zhao, R.; Xiao, T.; Wang, X. DeepReID: Deep Filter Pairing Neural Network for Person Re-Identification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014.

[18] Wei, L.; Zhang, S.; Gao, W.; Tian, Q. Person Transfer GAN to Bridge Domain Gap for Person Re-Identification. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 79–88.

[19] Liu, X.; Liu, W.; Mei, T.; Ma, H. A Deep Learning-Based Approach to Progressive Vehicle Re-identification for Urban Surveillance. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; pp. 869–884.

[20] Liu, H.; Tian, Y.; Wang, Y.; Pang, L.; Huang, T. Deep Relative Distance Learning: Tell the Difference between Similar Vehicles. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2167–2175.

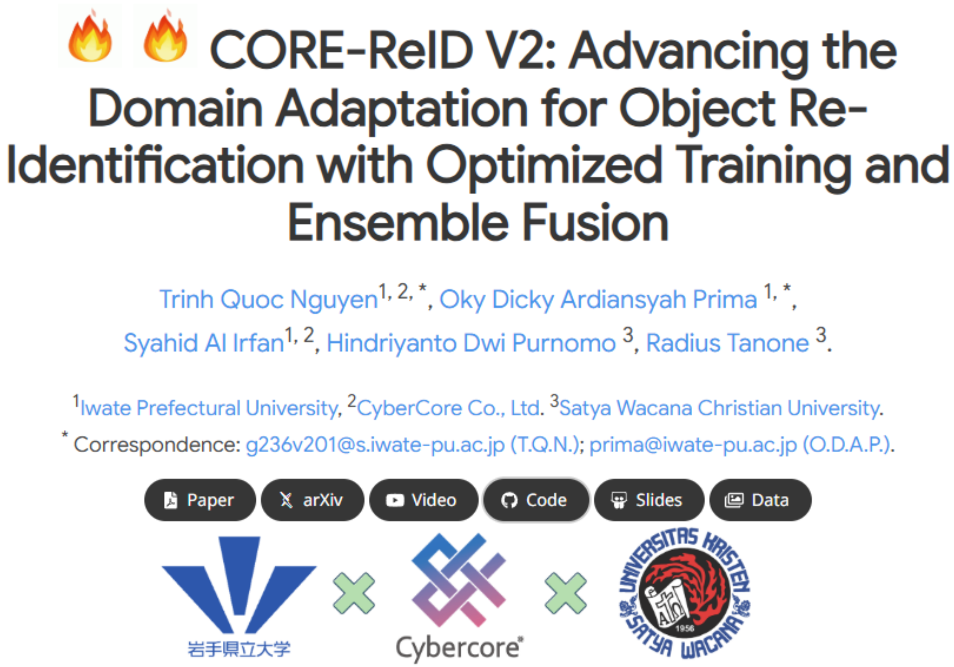
References

[21] Lou, Y.; Bai, Y.; Liu, J.; Wang, S.; Duan, L. VERI-Wild: A Large Dataset and a New Method for Vehicle Re-Identification in the Wild. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 3230–3238.

[22] Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.

CORE-ReID V2 Resources

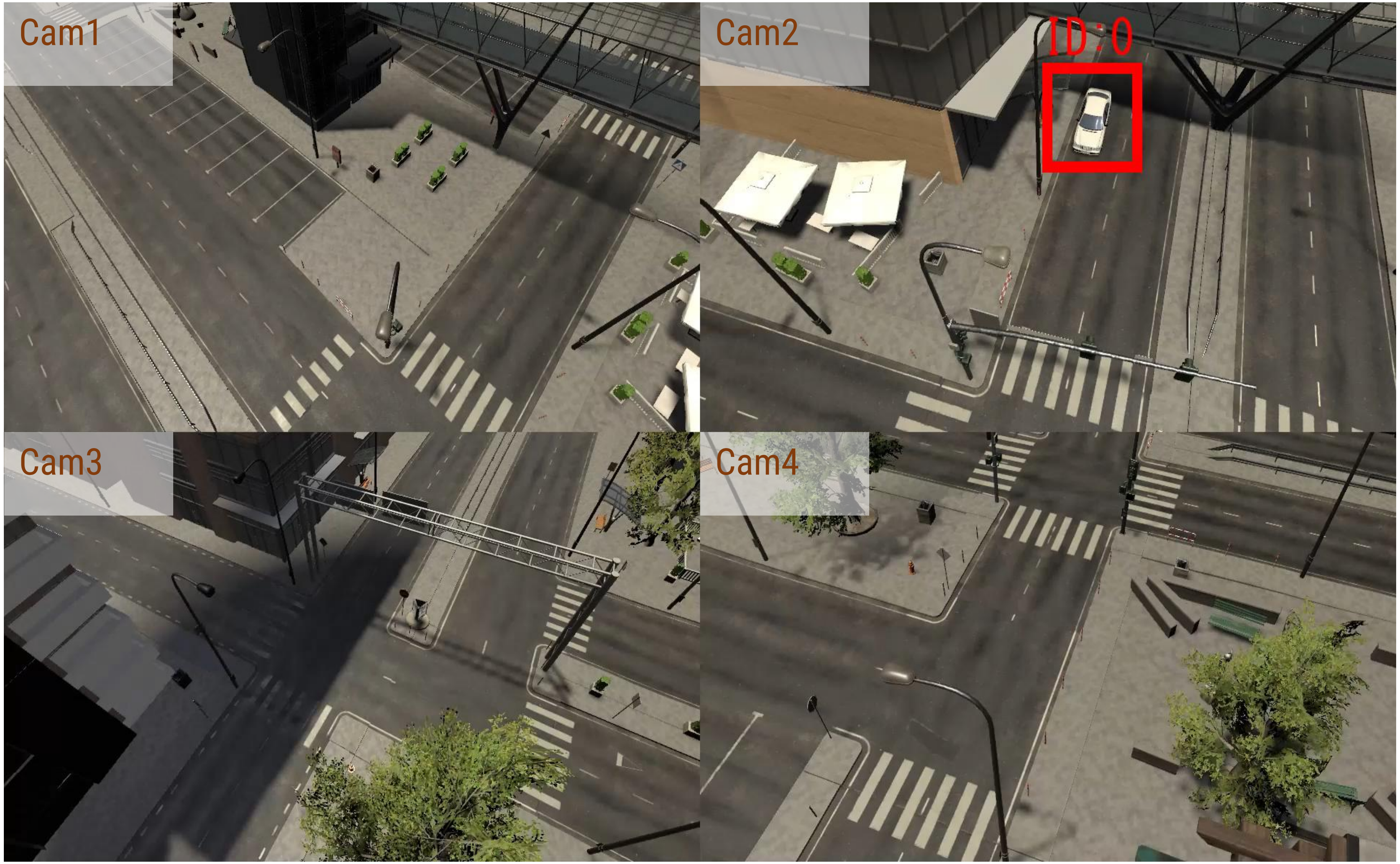
- Project Page: https://trinhquocnguyen.github.io/core-reid-v2-homepage/
- Github: https://github.com/TrinhQuocNguyen/CORE-ReID-V2
- Paper: https://www.mdpi.com/3042-5999/1/1/4
- Paperwithcode: https://paperswithcode.com/paper/core-reid-v2-advancing-the-domain-adaptation



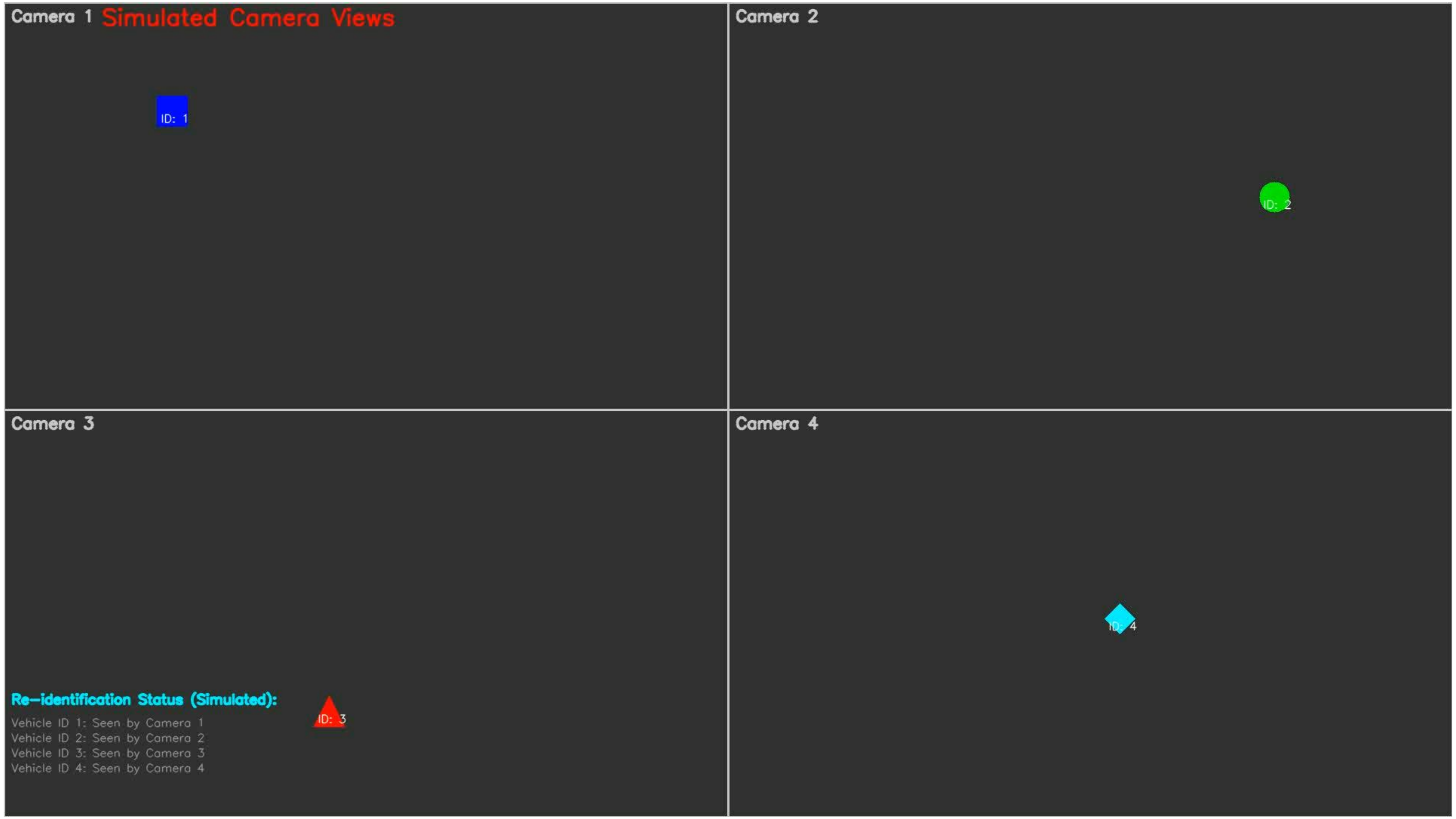
Ablation Study:: Demo Video (Person ReID)

Iwate Prefectural University

Ablation Study:: Demo Video (Vehicle ReID)



Ablation Study:: Demo Video (Vehicle ReID)



Iwate Prefectural University

Thank you