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

CORE-ReID: Comprehensive Optimization and Refinement through Ensemble Fusion in Domain Adaptation for Person Re-IDentification

CORE-ReID: ドメイン適応に基づく人物再認識のための大域特徴及び局所特徴のアンサンブル融合

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

2025/02/10

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

Tracking individuals across multiple camera views presents challenges that traditional tracking algorithms often fail to address.

Motivations

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)

Solution

Person Re-Identification

Research Background

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~)

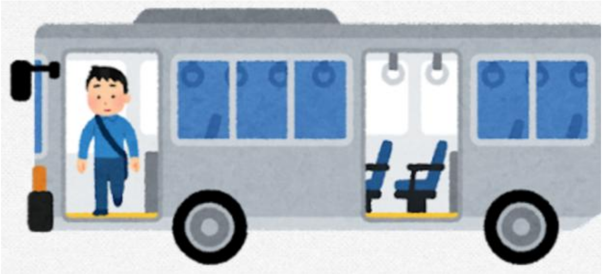
Source: https://morioka-machidukuri.jp/

Human Flow Analysis at Kochi City from (2024~)

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

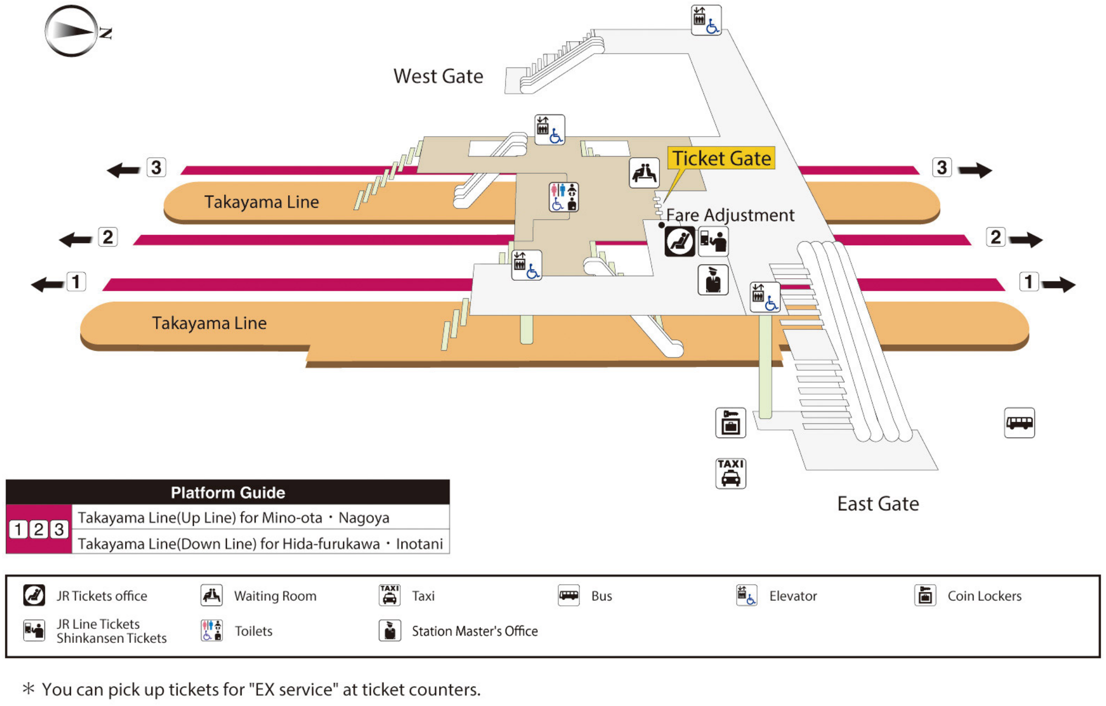
Origin-Destination Survey

Get on the bus/train



Get off the bus/train

Integration with edge devices

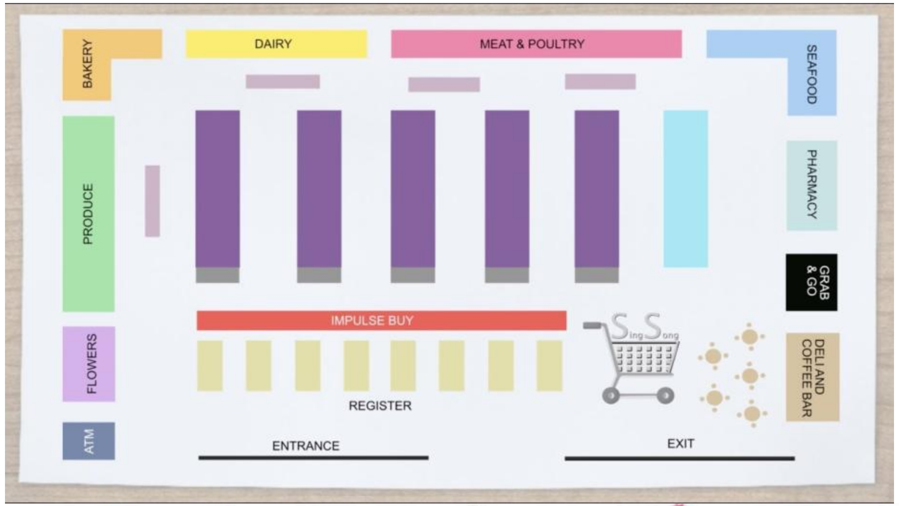


Analysis within the station platform

Research Background

Behavior Analysis

Shopping behavior



Research Background

Person Re-Identification

Query Image

Image Gallery

Research Background

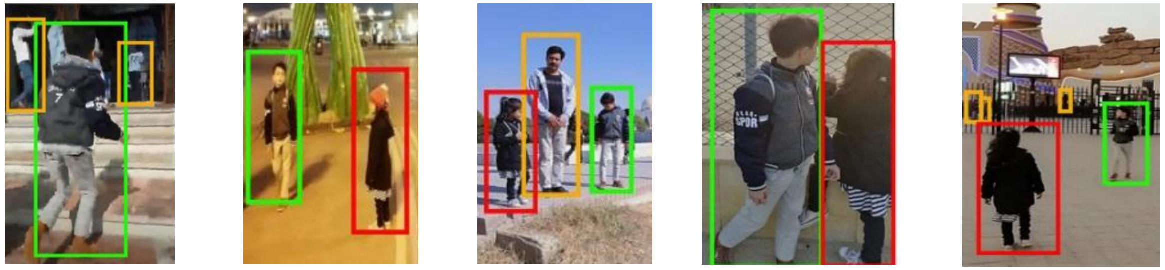
Person Re-Identification

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.

Camera

Person Re-Identification

Query Images

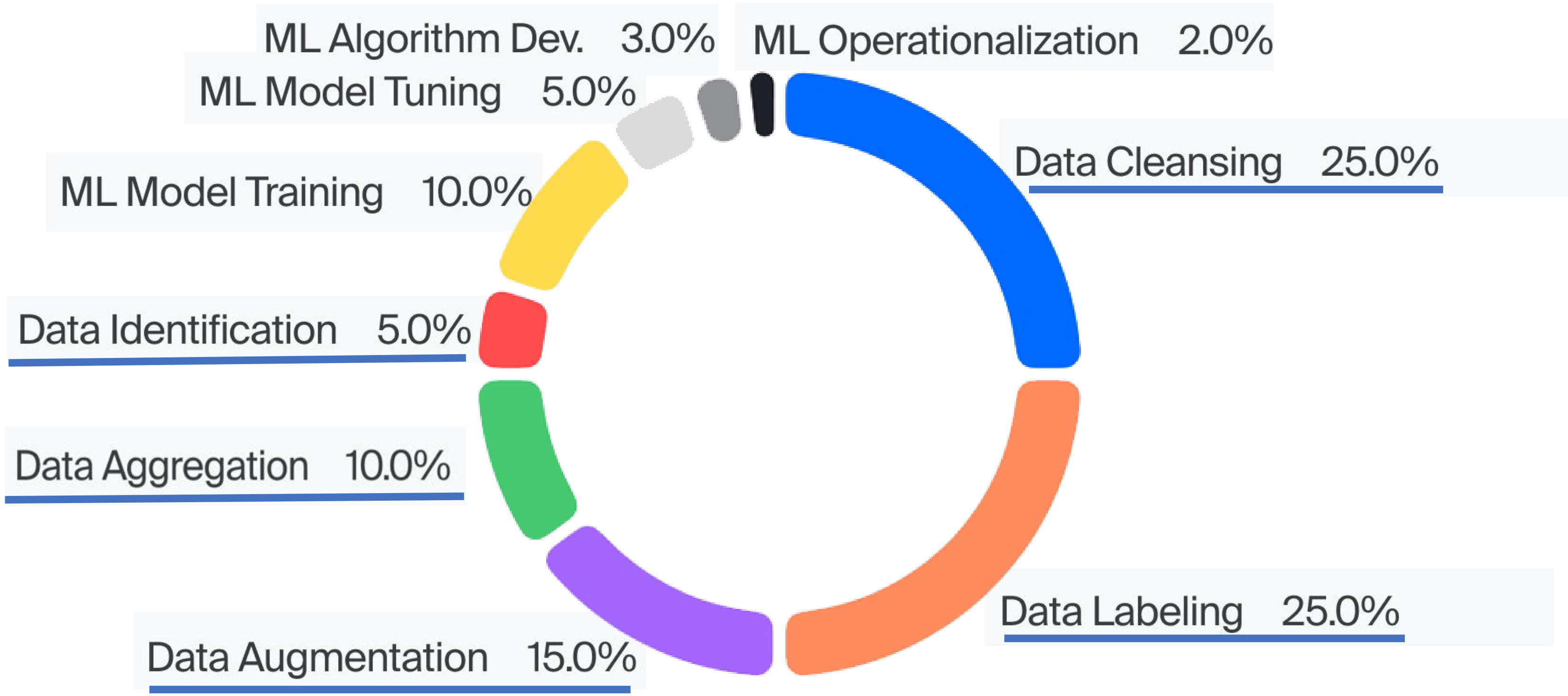


Camera #1 Camera #2 Camera #3 Camera #4 Camera #5

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

Iwate Prefectural University

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)

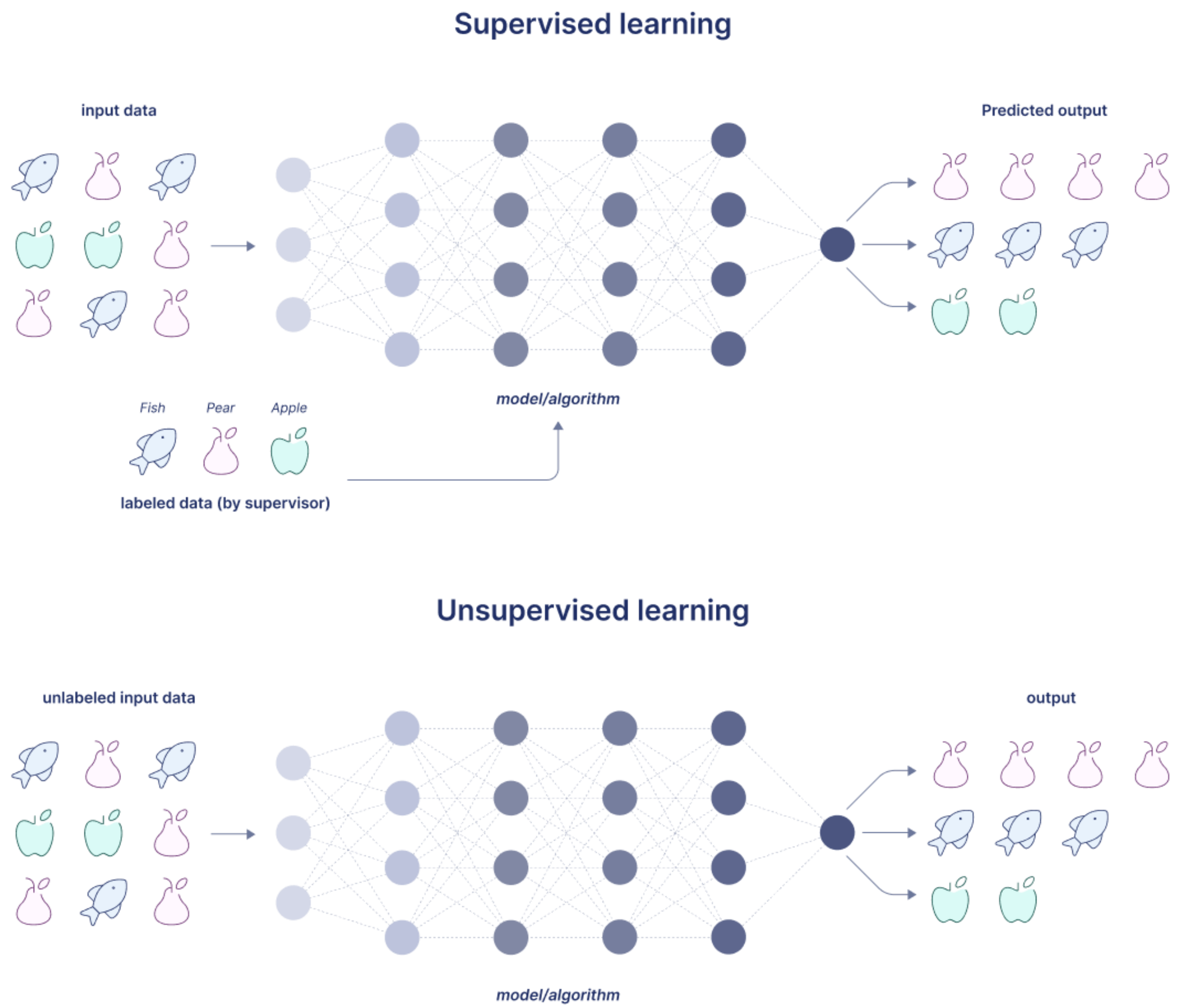
Related Work

Various Person Re-Identification Methods

- Person Re-Identification (ReID)
  - Supervised ReID
    - Pose-guided ReID
    - Attribute-guided ReID
    - Spatial-temporal ReID
    - …
    - Time-intensive tasks required for image labeling.
  - Unsupervised ReID
    - Direct Transfer
    - Domain Adaptation
      - Adaptation of the pre-trained model is required to achieve optimal performance on the new domain or dataset.
  - Generalizable ReID
    - Trained on a source dataset and directly evaluated on a target dataset.
  - Lifelong ReID
    - Continual Learning: learn continuously from new data without forgetting previously learned knowledge.
  - Others
    - Cross-modality ReID
    - Long-term ReID
    - Group ReID
    - Person Search
    - …

Related Work

Supervised vs Unsupervised Learning



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

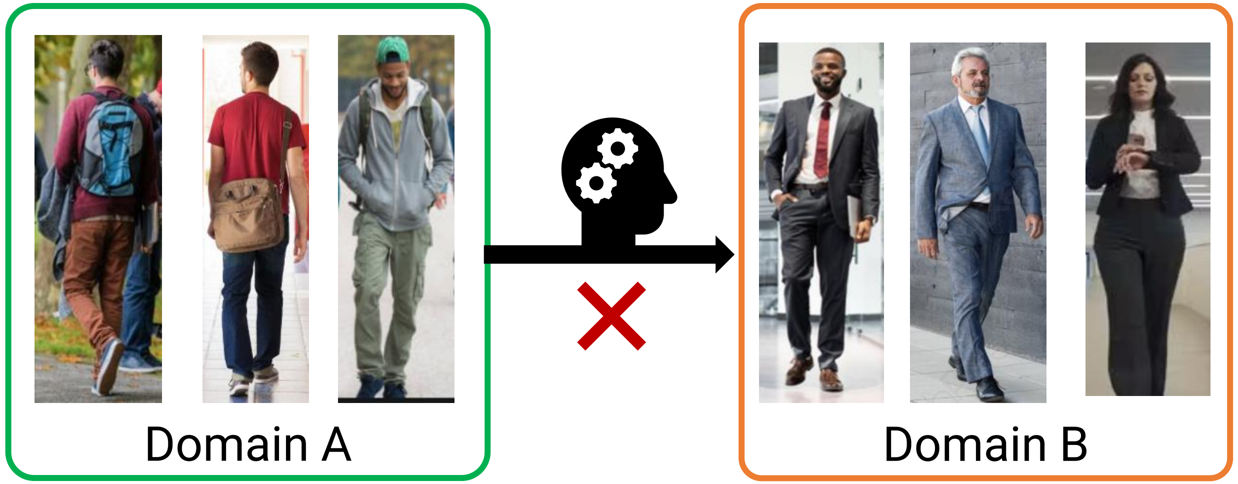
- Struggles with domain adaptation because it relies on labeled data from the target domain.
- Requires additional labeling efforts when adapting to a new domain.

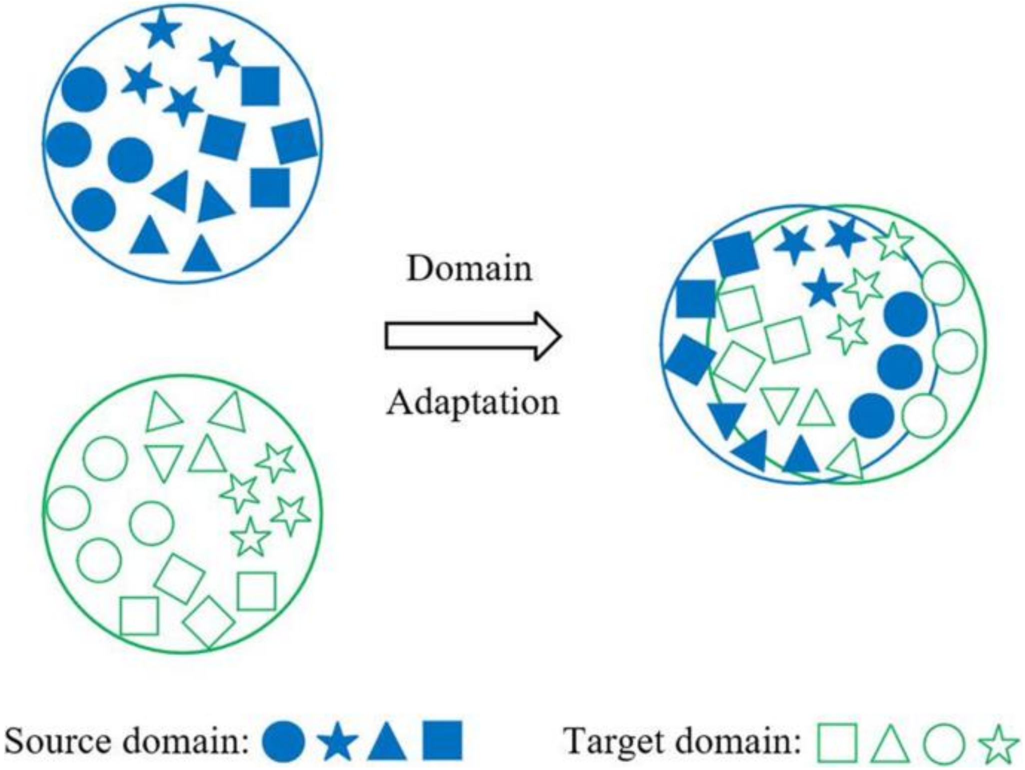
Unsupervised learning, while less accurate, is more flexible, scalable, and better suited for real-world applications where data annotation is impractical.

- 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.

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

Cross-Domain Adaption





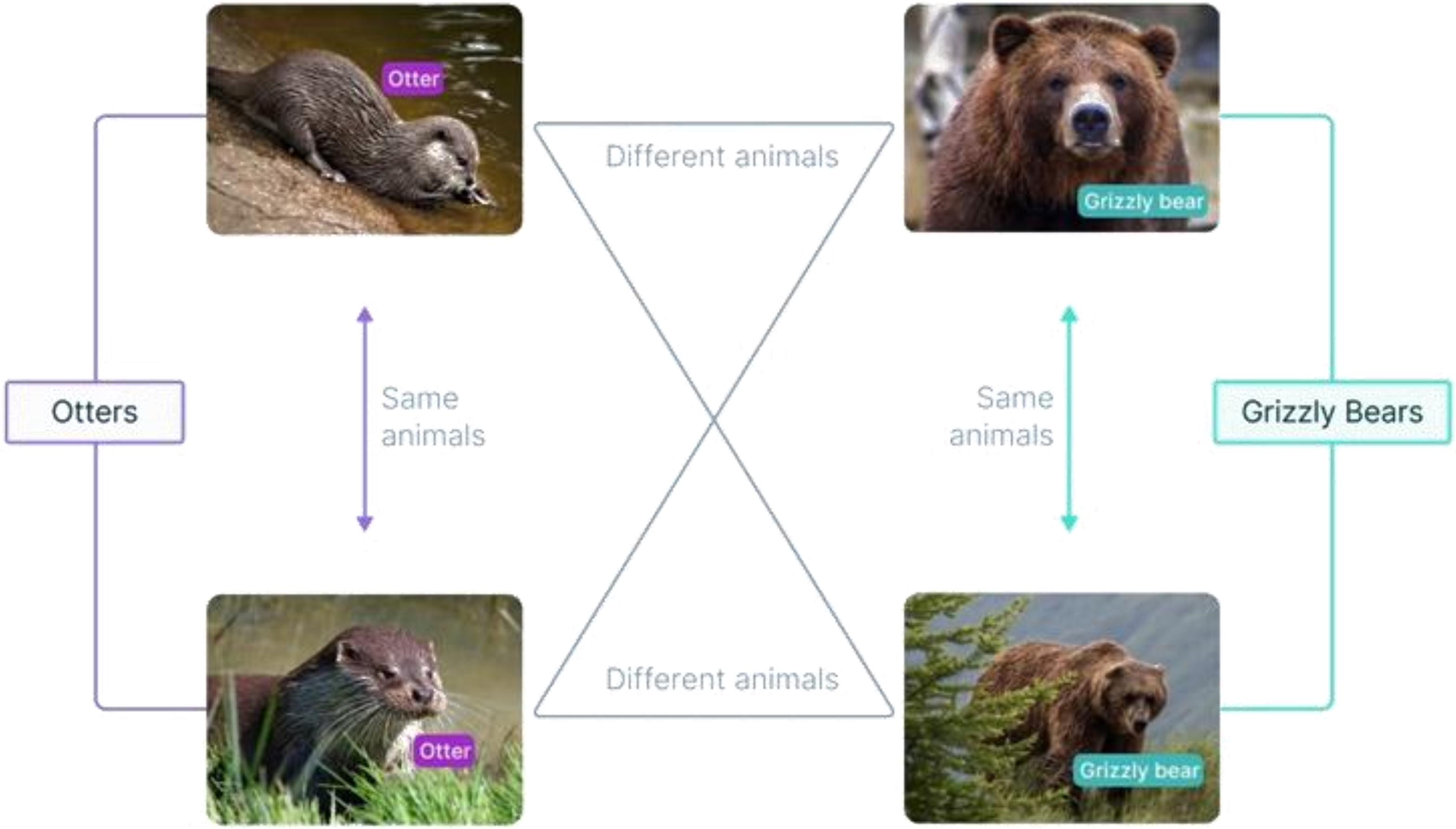
Re-identification (Re-ID) algorithms often struggle to generalize effectively across different domains.

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

Related Work

- Effective Model Training: Contrastive learning

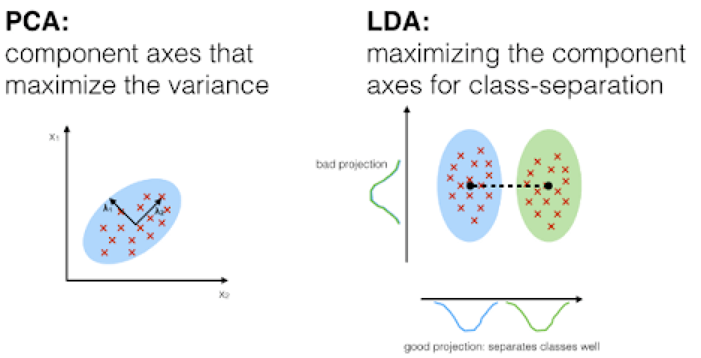
Contrastive learning extracts meaningful representations by distinguishing between positive and negative instance pairs.

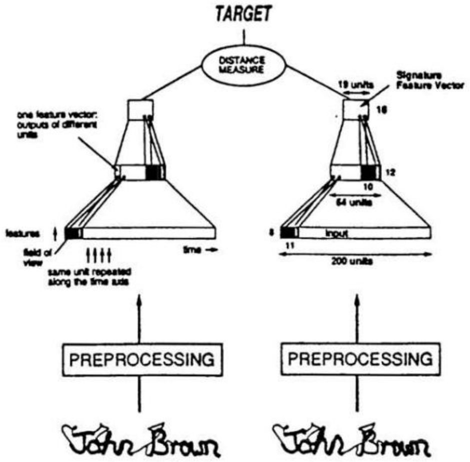


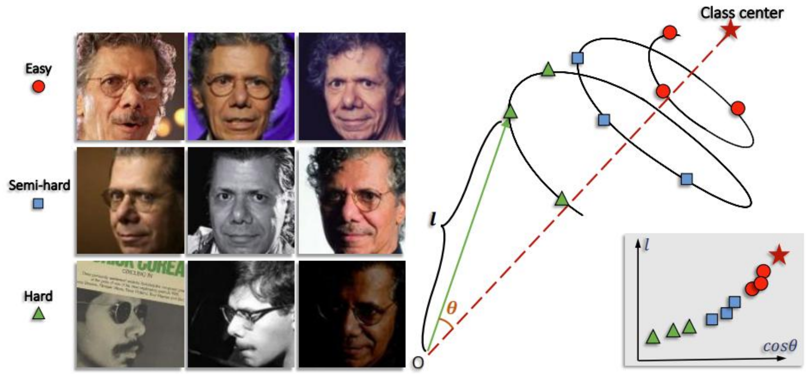
Related Work

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







Contrastive learning (Deep Learning Era 2010s)

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

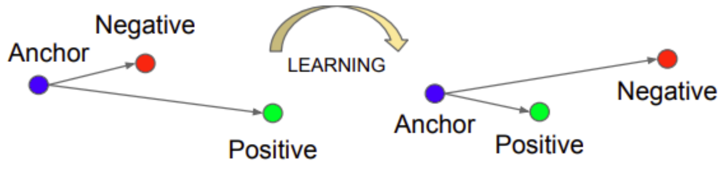
Supervised Contrastive Learning (2017-2019) [6]

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

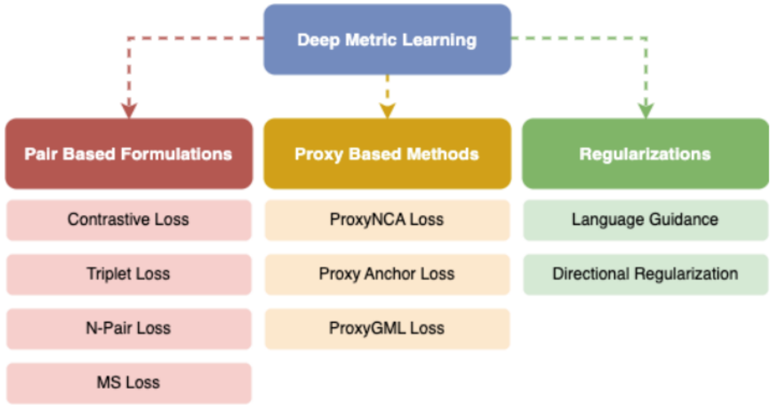
The **FaceNet** model (Schroff et al., 2015) leveraged **triplet loss**, a more advanced contrastive learning loss function, for face verification.

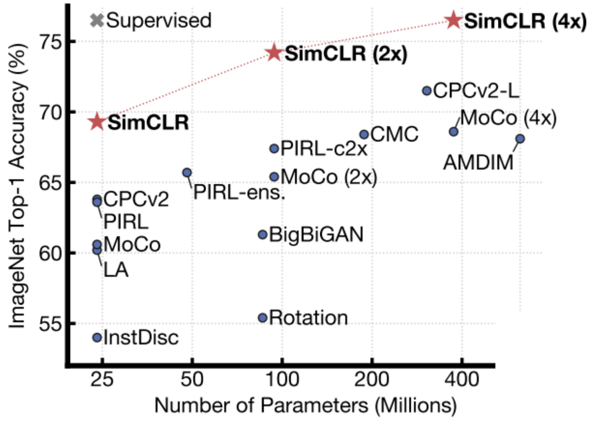
Deep Metric Learning and **Supervised Contrastive Learning** extended these ideas.

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



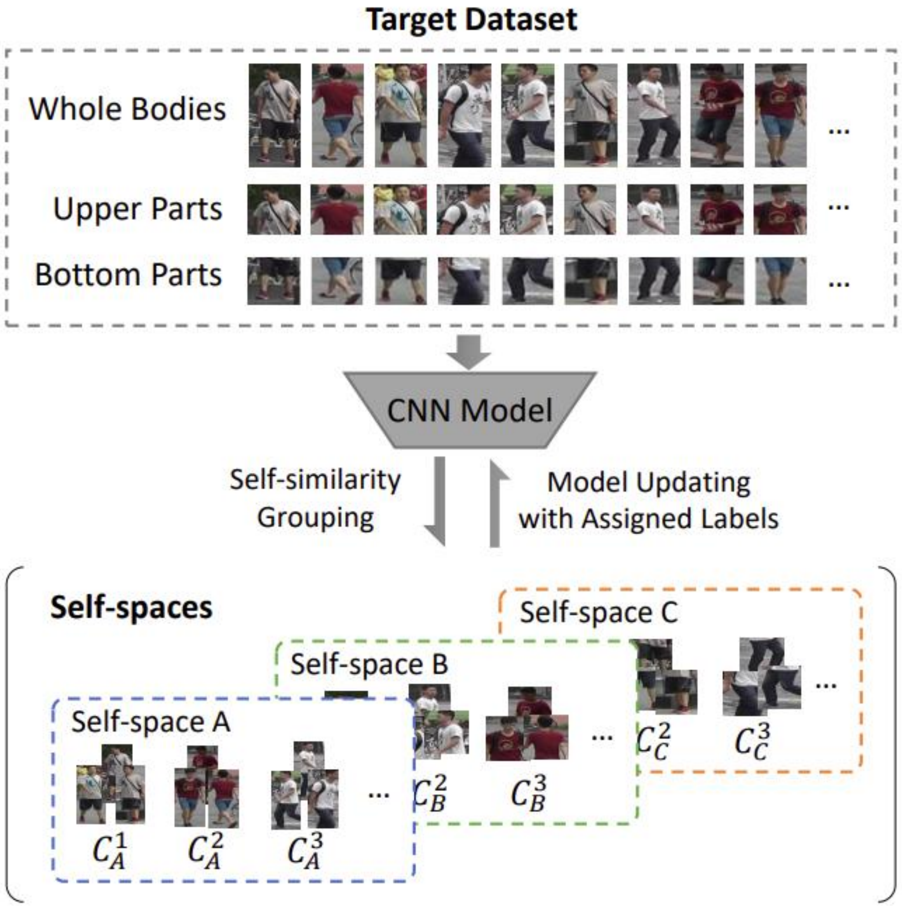
The Triplet Loss minimizes the distance between an anchor and a positive





Self-similarity Grouping (SSG) [12]

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



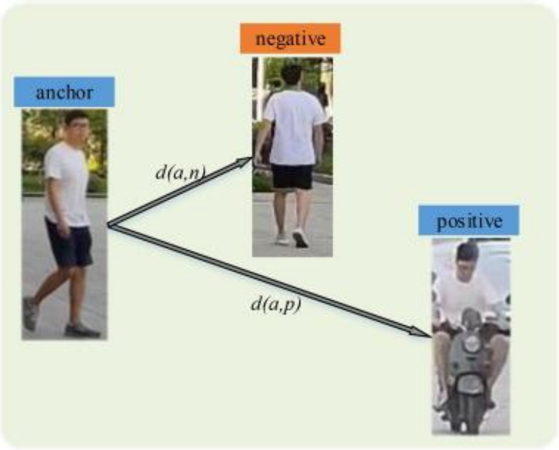
① SSG uses a single network for feature extraction in clustering, which is susceptible to the generation of numerous noisy pseudo-labels.

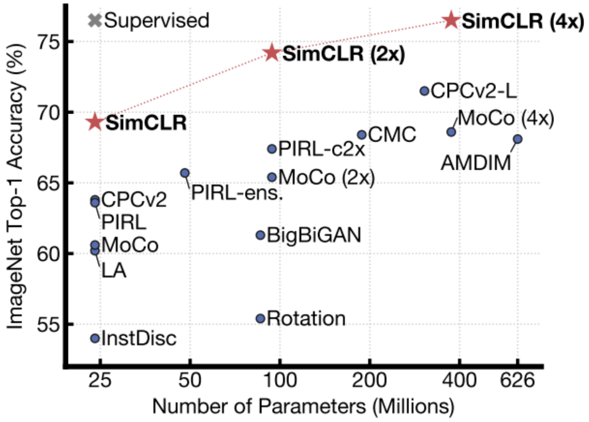
② SSG performs clustering based on global and local features independently, resulting in unlabeled samples acquiring multiple different pseudo-labels, leading to ambiguity in identity classification during training.

Related Work

Summary of related work

Early Works (2016-2018)	Modern Contrastive Learning in ReID (2019-Present)	This study: • Contrastive Learning • Ensemble (local & global features) • ReID
• 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



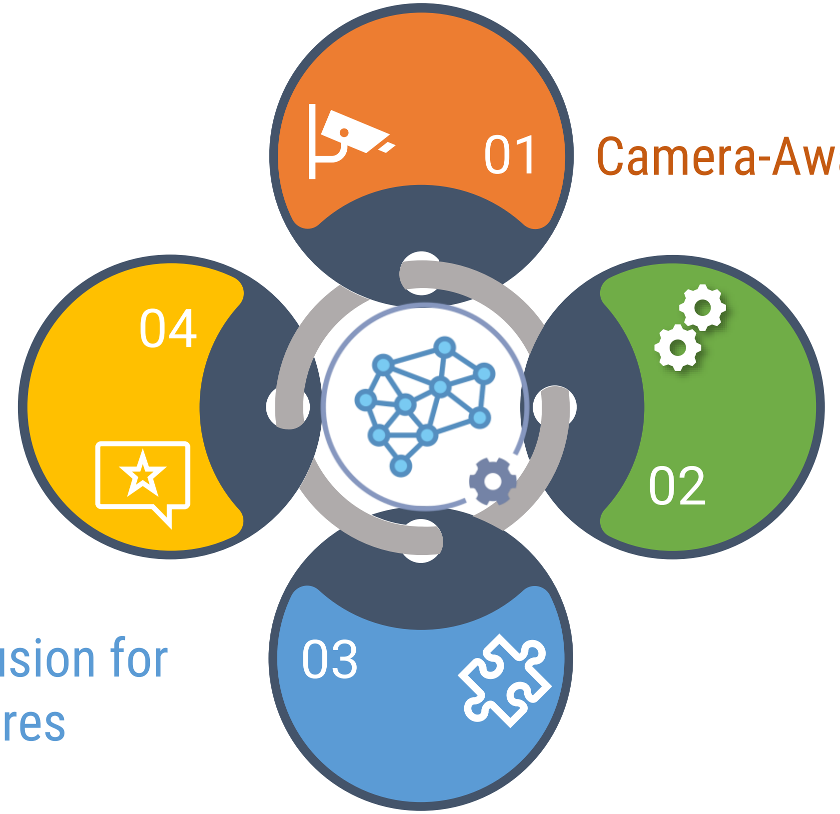


We are here

Iwate Prefectural University

Research Aim

This study aims to tackle the Unsupervised Domain Adaptation (UDA) problem in Person Re-identification by introducing a novel framework that optimizes and refines the adaptation process through the Ensemble Fusion component.



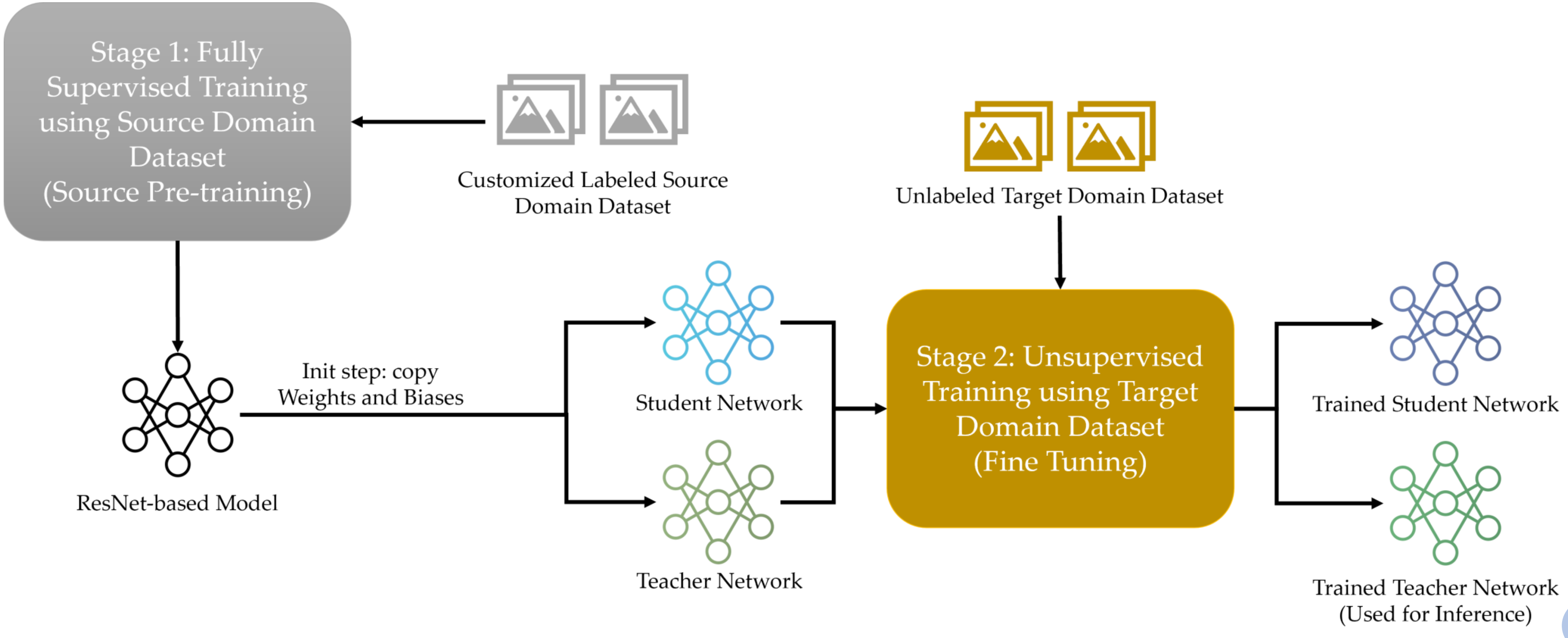
Proposed Methods

- Person Re-Identification (ReID)
  - Supervised ReID
    - Pose-guided ReID
    - Attribute-guided ReID
    - Spatial-temporal ReID
    - …
  - Unsupervised ReID
    - Direct Transfer
    - Domain Adaptation
      - (1) Style-transferred source-domain images
      - (2) Clustering-based methods
      - (3) Feature alignment
  - Generalizable ReID
  - Lifelong ReID
  - Others
    - Cross-modality ReID
    - Long-term ReID
    - Group ReID
    - Person Search
    - …

Iwate Prefectural University

Methodology

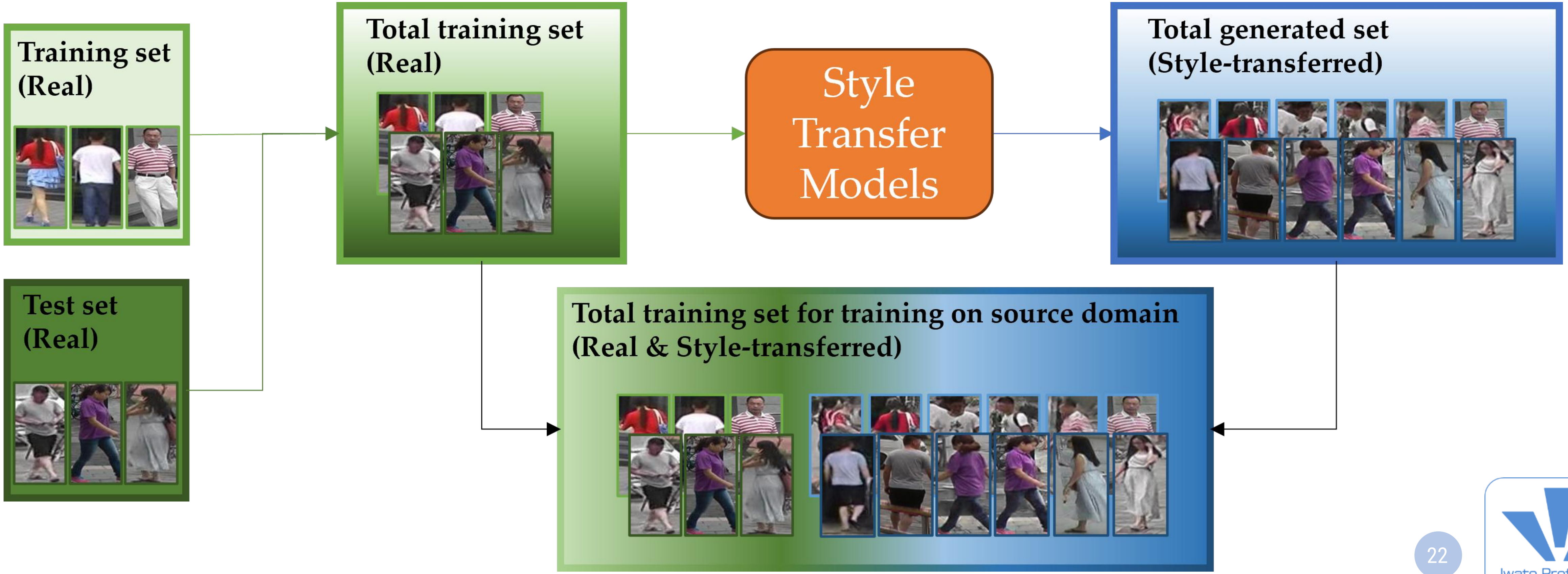
There are two stages: pre-training the model on the source domain in a fully supervised manner and fine-tuning the model on the target domain using an unsupervised learning approach.



Methodology:: Pre-Training

- Pre-Training: Camera-aware Image-to-Image translation on source dataset

Create the full training set for the source domain



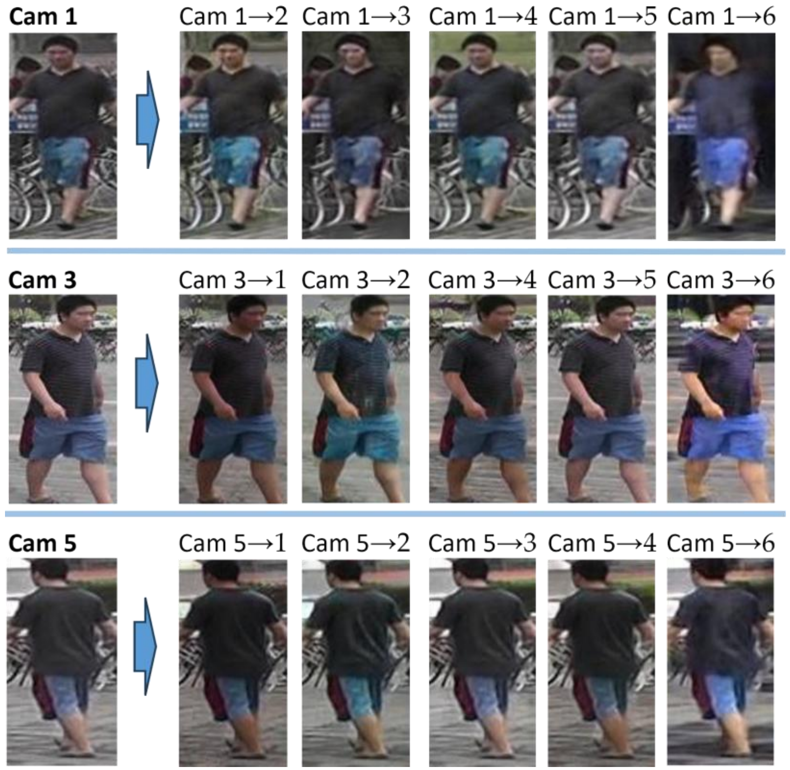
Iwate Prefectural University

Methodology:: Pre-Training

Pre-Training: Camera-aware Image-to-Image translation on source dataset

CycleGAN [13] was used to build Style Transfer Models

Training data from Market-1501 dataset



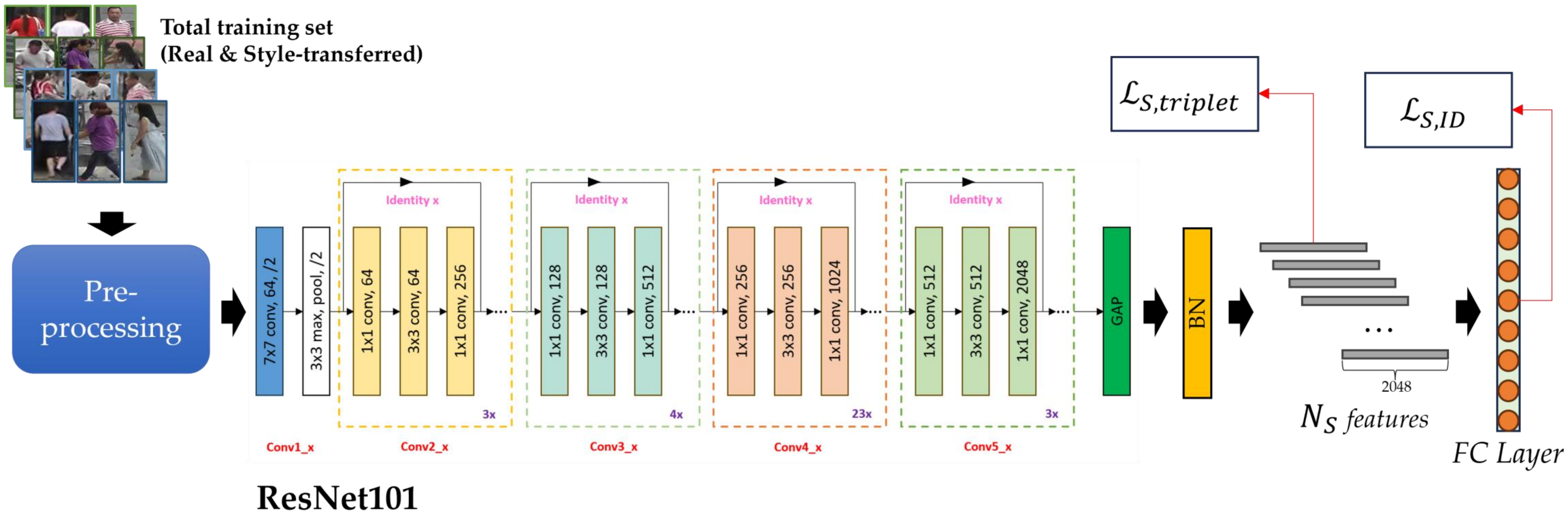
Test data from Market-1501 dataset

Iwate Prefectural University

Methodology:: Pre-Training

Pre-Training: Source-domain pre-training

Supervised Pre-Training



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

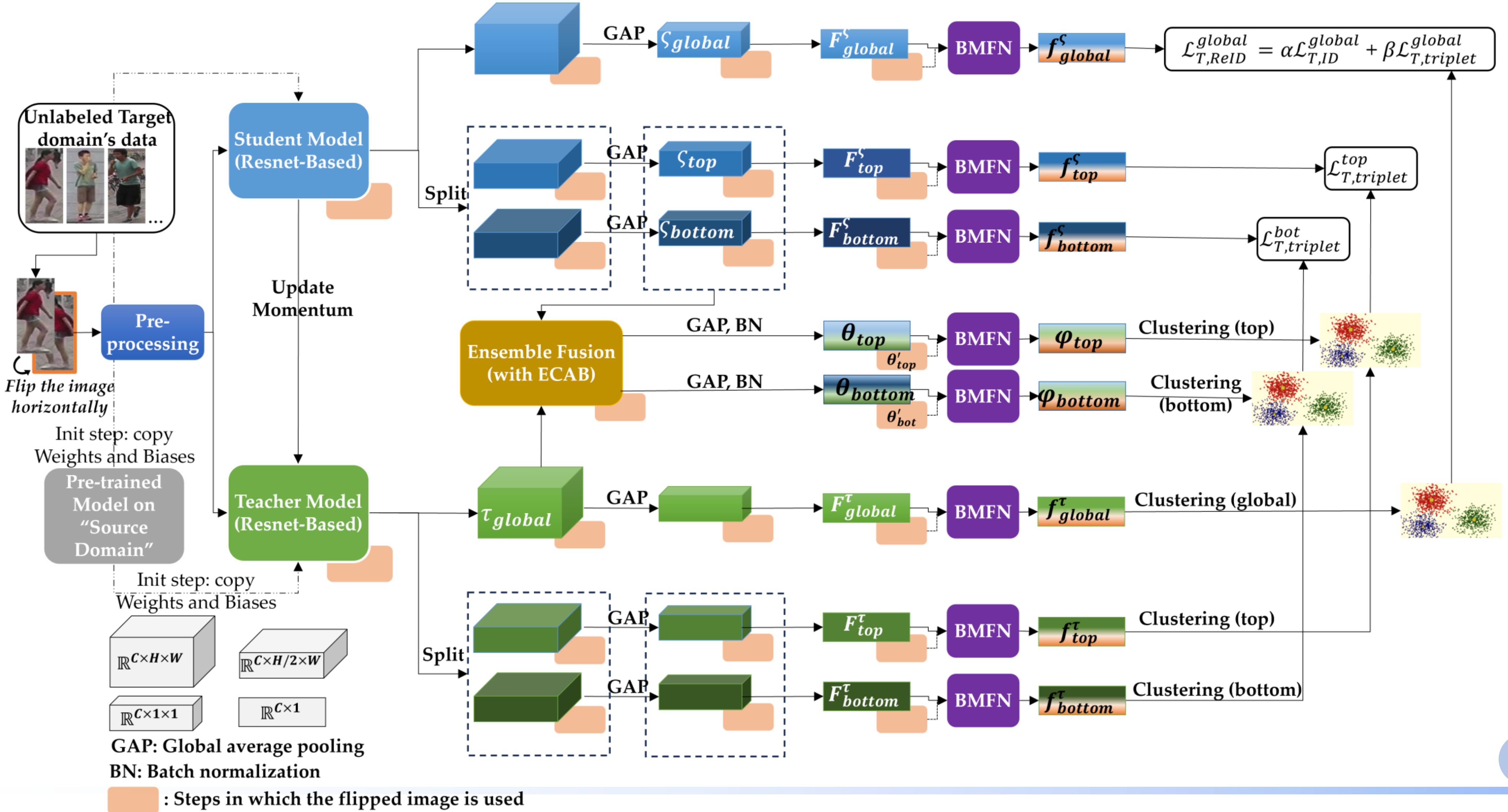
Methodology:: Fine Tuning

Fine Tuning: Target-domain fine-tuning

Needs	Related work	This study
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)

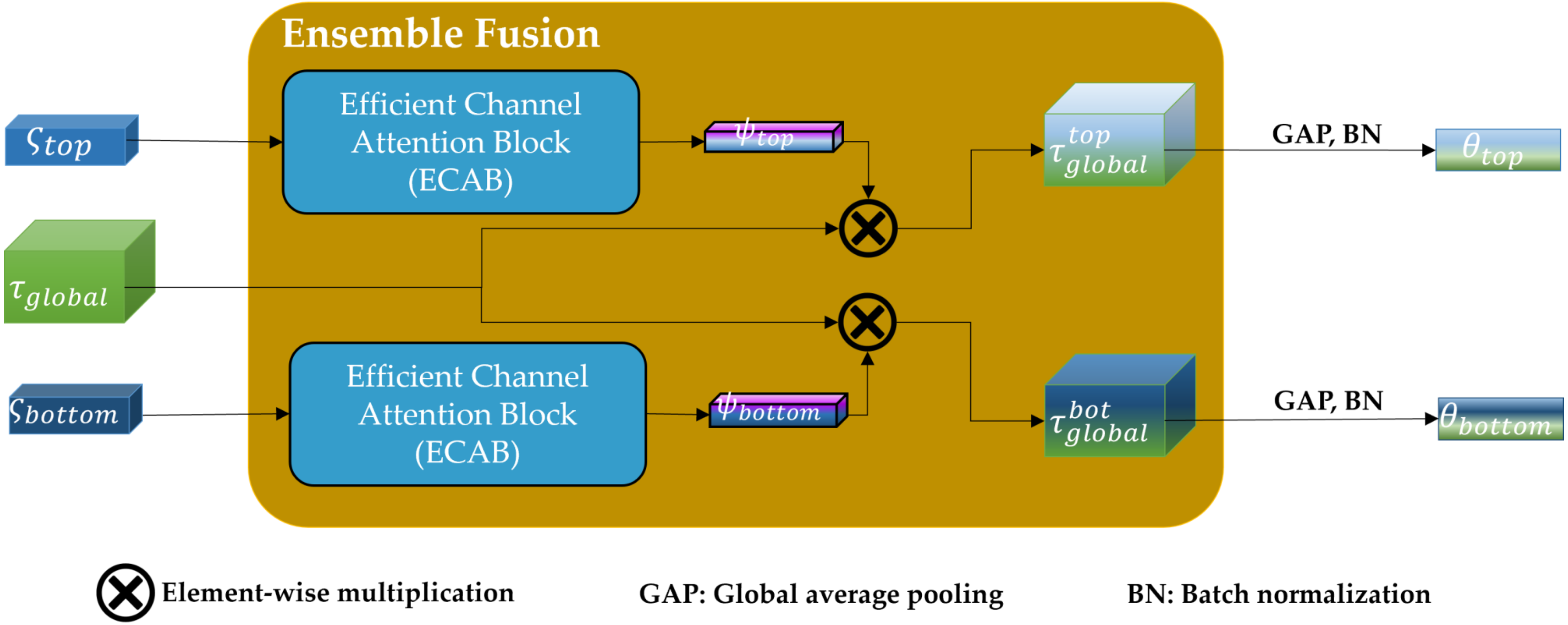
Methodology:: Fine Tuning

Fine Tuning: Target-domain fine-tuning



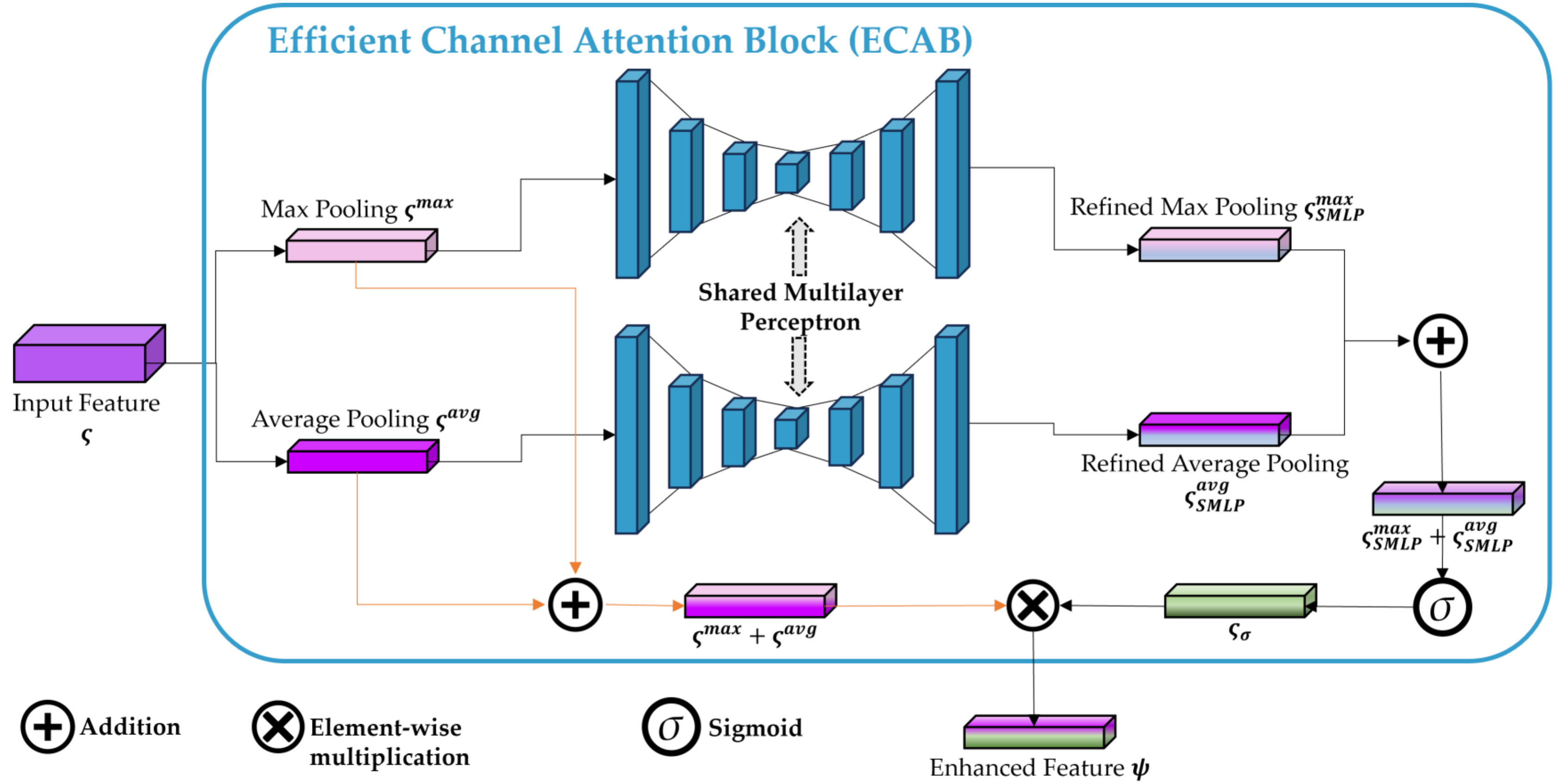
Fine Tuning: Target-domain fine-tuning

Ensemble Fusion: combine the Global and Local (Top and Bottom) features



Fine Tuning: Target-domain fine-tuning

The Efficient Channel Attention Block (ECAB) enhances representation capability by employing attention mechanisms that prioritize critical features while suppressing redundant ones.



Methodology:: Fine Tuning

Fine Tuning: Target-domain fine-tuning

Bidirectional Mean Feature Normalization (BMFN) involving both original and horizontally flipped vectors

Given an image $x_{T,i}$ in target domain dataset, and its flipped image $x'_{T,i}$. After getting the feature map F_j^m and its paired flipped image's feature map F'^m_j, j ∈ $\{global, top, bottom\}, m \in \{\varsigma, \tau\}$. The outputs from BMFN can be calculated as:

$$f_j^m = \frac{\frac{F_j^m + F'^m_j}{2}}{\left\| \frac{F_j^m + F'^m_j}{2} \right\|_2}$$

Original Image

Horizontally Flipped Image

Iwate Prefectural University

Results:: Evaluation Datasets

- Three benchmark datasets

Dataset	Cameras	Training Set (ID/Image)	Test Set (ID/Image)	
			Gallery	Query
Market-1501	6	751/12,936	750/19,732	750/3368
CUHK03	2	767/7365	700/5332	700/1400
MSMT17	15	1401/32,621	3060/82,161	3060/11,659

Market-1501 [16]

CUHK03 [17]

MSMT17 [18]

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] [50]	CVPR 2020	17.5	17.1	-	-	52.4	77.8	-	-
UDAR [51]	PR 2020	20.9	20.3	-	-	56.6	77.1	-	-
QAConv$_{50}$ [a] [52]	ECCV 2020	32.9	33.3	-	-	66.5	85.0	-	-
M^{3}L [a] [53]	CVPR 2021	35.7	36.5	-	-	62.4	82.7	-	-
MetaBIN [a] [54]	CVPR 2021	43.0	43.1	-	-	67.2	84.5	-	-
DFH-Baseline [55]	CVPR 2022	10.2	11.2	-	-	13.2	31.1	-	-
DFH [a] [55]	CVPR 2022	27.2	30.5	-	-	31.3	56.5	-	-
META [a] [56]	ECCV 2022	47.1	46.2	-	-	76.5	90.5	-	-
ACL [a] [57]	ECCV 2022	49.4	50.1	-	-	76.8	90.6	-	-
RCFA [58]	Electronics 2023	17.7	18.5	33.6	43.4	34.5	63.3	78.8	83.9
CRS [59]	JSJTU 2023	-	-	-	-	65.3	82.5	93.0	95.9
MTI [60]	JVCIR 2024	16.3	16.2	-	-	-	-	-	-
PAOA+ [a] [61]	WACV 2024	50.3	50.9	-	-	77.9	91.4	-	-
Baseline	Ours	<u>55.2</u>	<u>55.7</u>	<u>72.1</u>	<u>81.0</u>	<u>82.2</u>	<u>92.0</u>	<u>96.7</u>	<u>97.6</u>
CORE-ReID	Ours	**62.9**	**61.0**	**79.6**	**87.2**	**83.6**	**93.6**	**97.3**	**98.7**

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

Results:: Market → MSMT & CUHK → MSMT

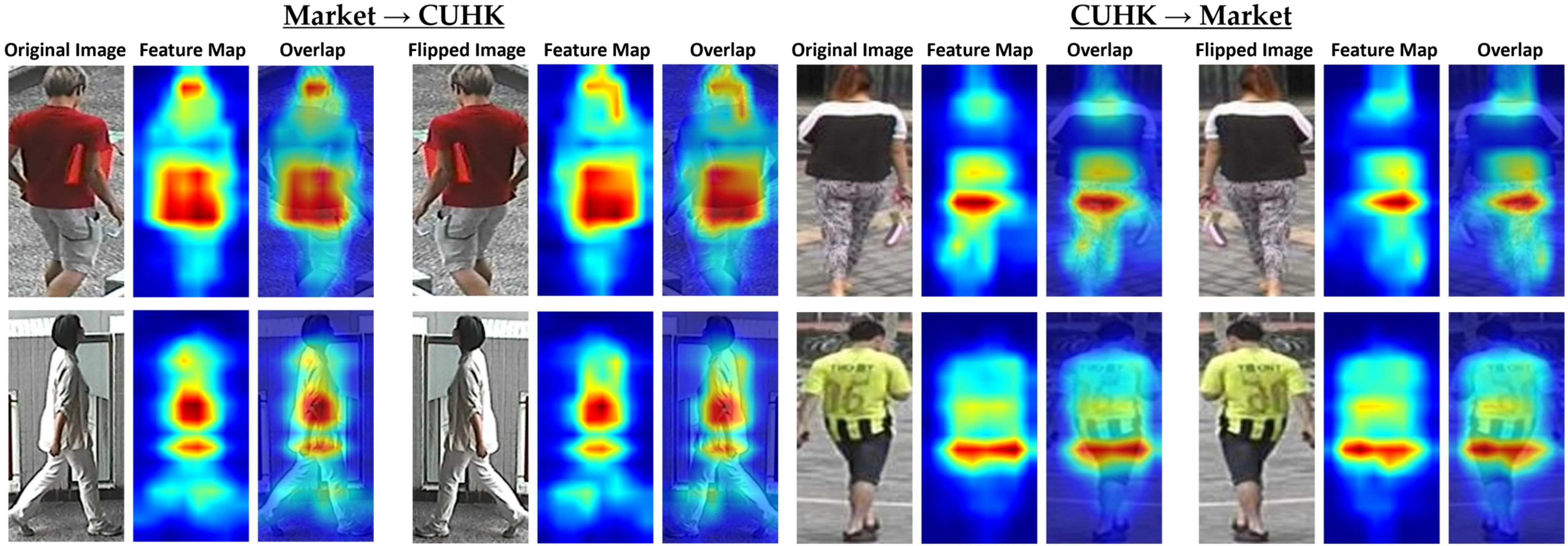
		Market → MSMT				CUHK → MSMT			
Method	**Reference**	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
NRMT [62]	ECCV 2020	19.8	43.7	56.5	62.2	-	-	-	-
DG-Net++ [38]	ECCV 2020	22.1	48.4	-	-	-	-	-	-
MMT [22]	ICLR 2020	22.9	52.5	-	-	13.5 [b]	30.9 [b]	44.4 [b]	51.1 [b]
UDAR [51]	PR 2020	12.0	30.5	-	-	11.3	29.6	-	-
Dual-Refinement [63]	arXiv 2020	25.1	53.3	66.1	71.5	-	-	-	-
SNR [a] [50]	CVPR 2020	-	-	-	-	7.7	22.0	-	-
$QAConv_{50}$ [a] [52]	ECCV 2020	-	-	-	-	17.6	46.6	-	-
M^3L [a] [53]	CVPR 2021	-	-	-	-	17.4	38.6	-	-
MetaBIN [a] [54]	CVPR 2021	-	-	-	-	18.8	41.2	-	-
RDSBN [64]	CVPR 2021	30.9	61.2	73.1	77.4	-	-	-	-
ClonedPerson [65]	CVPR 2022	14.6	41.0	-	-	13.4	42.3	-	-
META [a] [56]	ECCV 2022	-	-	-	-	24.4	52.1	-	-
ACL [a] [57]	ECCV 2022	-	-	-	-	21.7	47.3	-	-
CLM-Net [66]	NCA 2022	29.0	56.6	69.0	74.3	-	-	-	-
CRS [59]	JSJTU 2023	22.9	43.6	56.3	62.7	22.2	42.5	55.7	62.4
HDNet [67]	IJMLC 2023	25.9	53.4	66.4	72.1	-	-	-	-
DDNet [68]	AI 2023	28.5	59.3	72.1	76.8	-	-	-	-
CaCL [69]	ICCV 2023	36.5	66.6	75.3	80.1	-	-	-	-
PAOA+ [a] [61]	WACV 2024	-	-	-	-	26.0	52.8	-	-
OUDA [70]	WACV 2024	20.2	46.1	-	-	-	-	-	-
M-BDA [71]	VCIR 2024	26.7	51.4	64.3	68.7	-	-	-	-
UMDA [72]	VCIR 2024	32.7	62.4	72.7	78.4	-	-	-	-
Baseline	Ours	<u>40.1</u>	<u>67.3</u>	<u>79.4</u>	<u>83.1</u>	<u>37.2</u>	<u>65.5</u>	<u>77.2</u>	<u>81.0</u>
CORE-ReID	Ours	**41.9**	**69.5**	**80.3**	**84.4**	**40.4**	**67.3**	**79.0**	**83.1**

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.

Iwate Prefectural University

Ablation Study:: Feature Maps Visualization

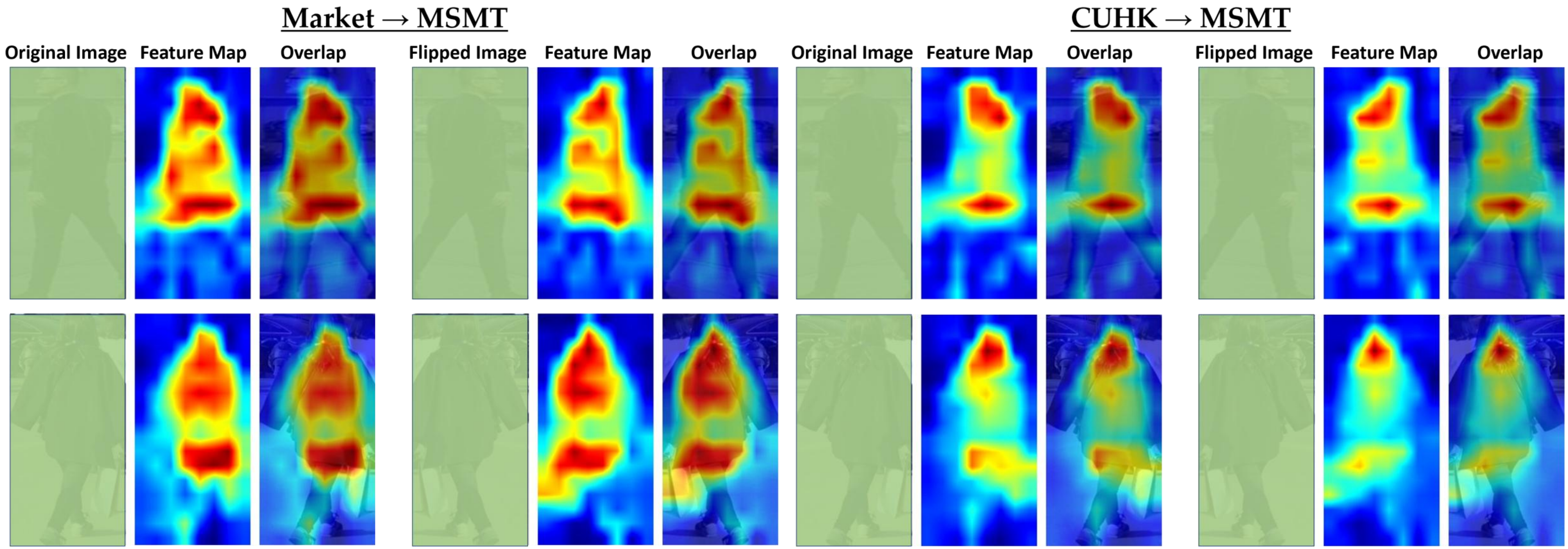
The feature map, visualized using Grad-CAM at the global feature level, highlights important features of each individual, represented as heatmaps.



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.



Iwate Prefectural University

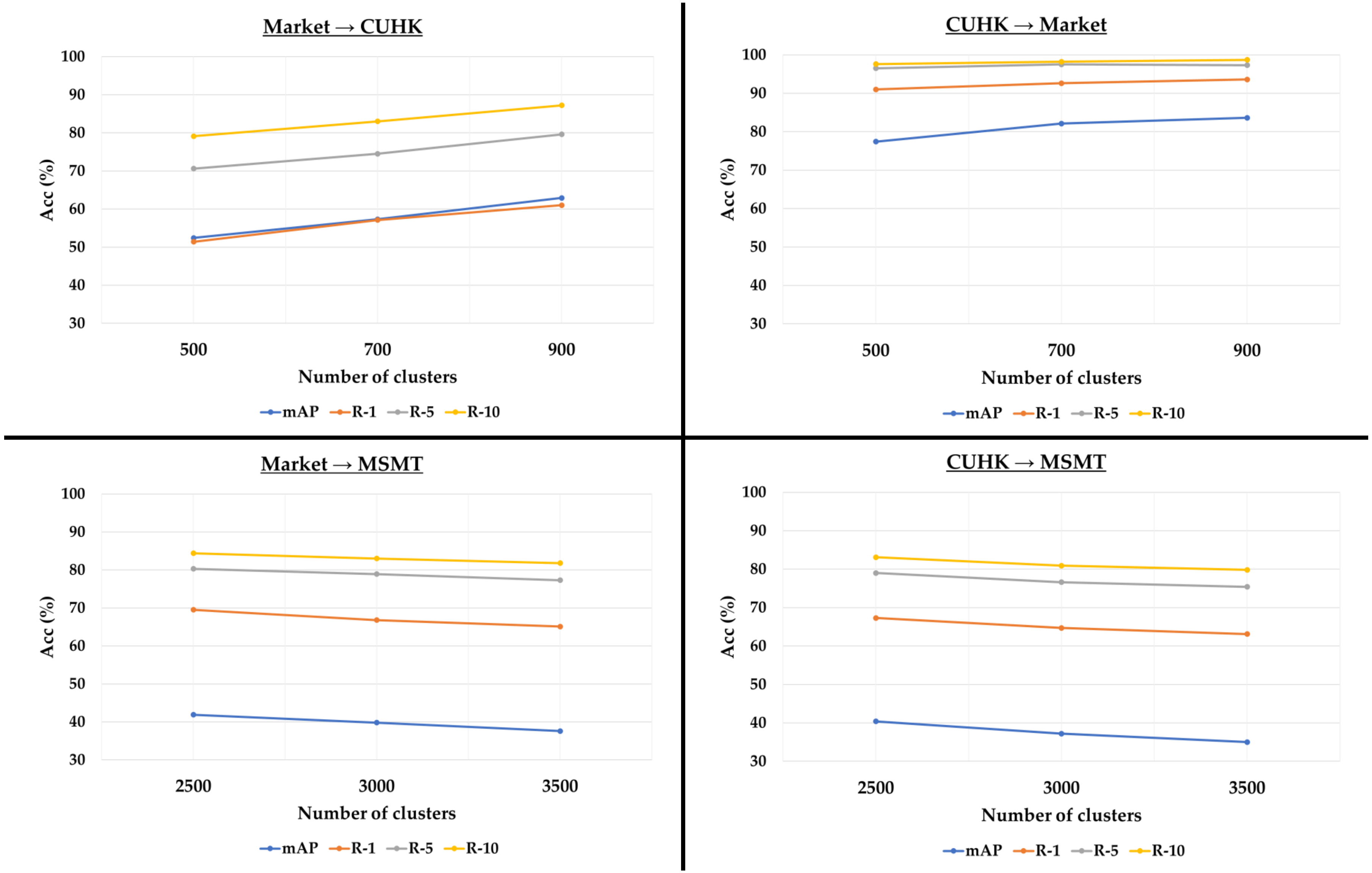
Ablation Study:: K-means Clustering Settings

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

	Market → CUHK				CUHK → Market			
Method	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 500$)	52.4	51.4	70.6	79.1	77.4	91.0	96.5	97.6
Ours ($M_{T,j} = 700$)	57.3	57.1	74.5	83.0	82.1	92.6	**97.5**	98.2
Ours ($M_{T,j} = 900$)	**62.9**	**61.0**	**79.6**	**87.2**	**83.6**	**93.6**	97.3	**98.7**
	Market → MSMT				CUHK → MSMT			
Method	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours ($M_{T,j} = 2500$)	**41.9**	**69.5**	**80.3**	**84.4**	**40.4**	**67.3**	**79.0**	**83.1**
Ours ($M_{T,j} = 3000$)	39.8	66.8	78.9	83.0	37.2	64.7	76.6	80.9
Ours ($M_{T,j} = 3500$)	37.6	65.1	77.3	81.8	35.0	63.1	75.4	79.8

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

Ablation Study:: K-means Clustering Settings



Ablation Study:: ECAB and BMFN Settings

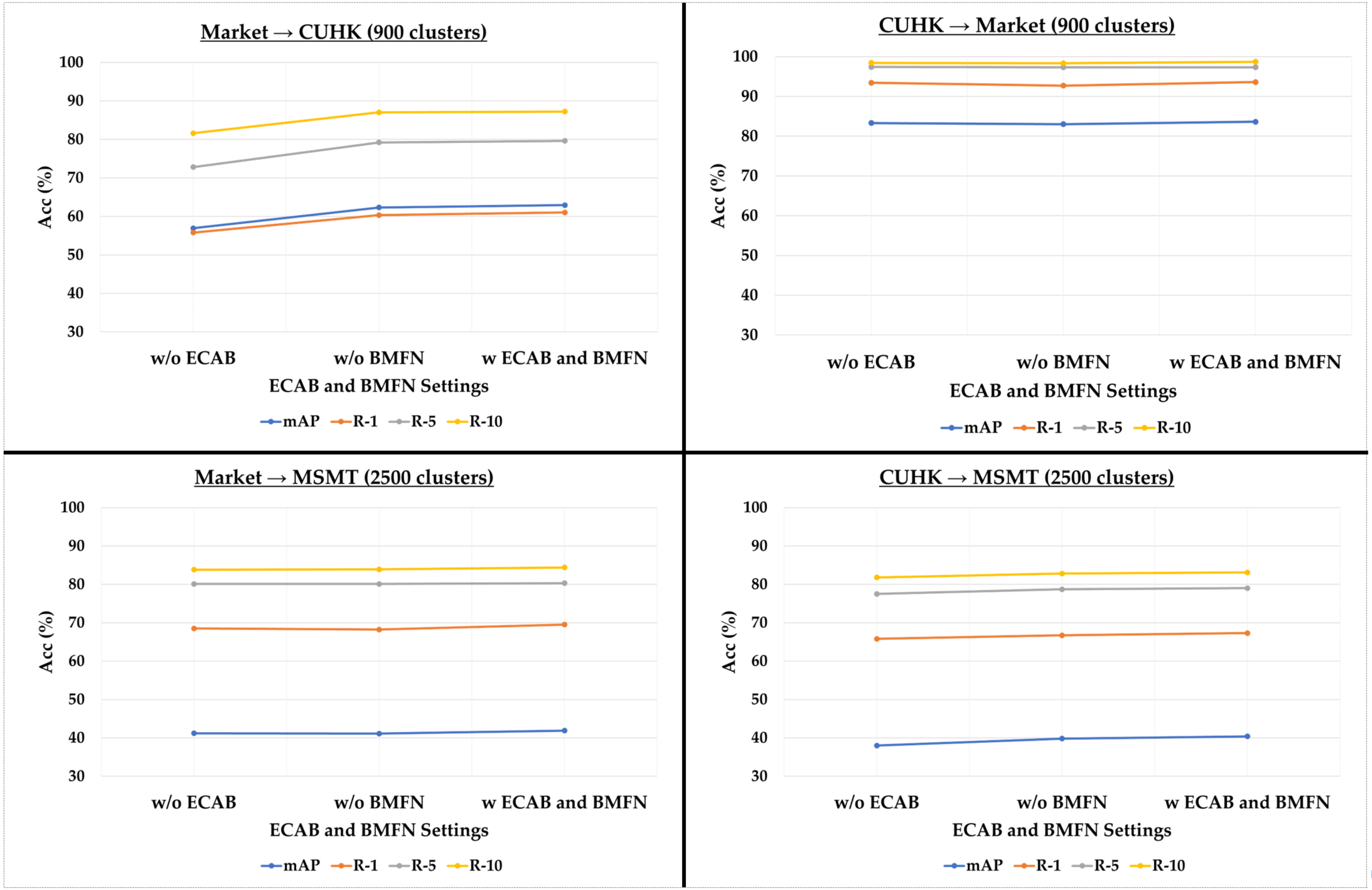
To validate the effectiveness of ECAB and BMFN, we performed an experiment where it is removed from our network.

	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 ECAB)	56.9	55.8	72.8	81.6	83.3	93.4	**97.4**	98.4
Ours (without BMFN)	62.3	60.3	79.2	87.0	83.0	92.7	97.3	98.3
Ours (with ECAB and BMFN)	**62.9**	**61.0**	**79.6**	**87.2**	**83.6**	**93.6**	97.3	**98.7**
	Market → MSMT ($M_{T,j} = 2500$)				**CUHK → MSMT ($M_{T,j} = 2500$)**			
Method	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
Ours (without ECAB)	41.2	68.5	80.1	83.8	38.0	65.8	77.5	81.8
Ours (without BMFN)	41.1	68.2	80.1	83.9	39.8	66.7	78.7	82.8
Ours (with ECAB and BMFN)	**41.9**	**69.5**	**80.3**	**84.4**	**40.4**	**67.3**	**79.0**	**83.1**

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

Iwate Prefectural University

Ablation Study:: ECAB and BMFN Settings



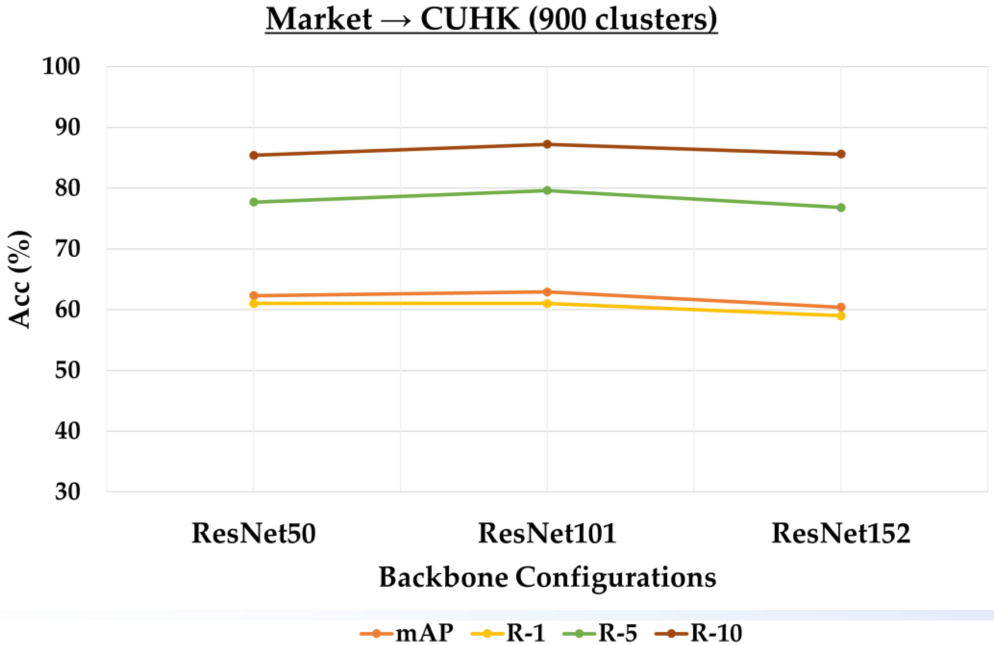
Iwate Prefectural University

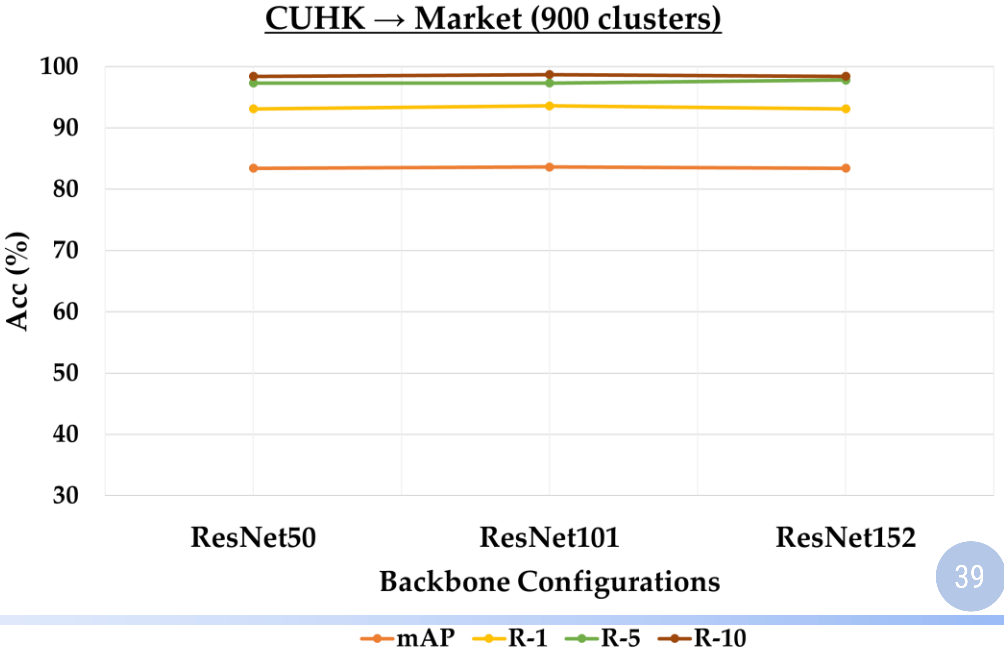
Ablation Study:: Backbone Configurations

The performance of different backbone architectures (ResNet50, ResNet101, and ResNet152)

	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 (ResNet50)	62.3	61.0	77.7	85.4	83.4	93.1	97.3	98.4
Ours (ResNet101)	**62.9**	**61.0**	**79.6**	**87.2**	**83.6**	**93.6**	97.3	**98.7**
Ours (ResNet152)	60.4	59.0	76.8	85.6	83.4	93.1	**97.8**	98.4

Bold denotes the best results.





Ablation Study:: Additional Experiments

		Market To Duke				Duke To Market			
Method	**Reference**	**mAP**	**mAP**	**mAP**	**mAP**	**mAP**	**R-1**	**R-5**	**R-10**
PDA-Net [16]	ICCV 2019	45.1	63.2	77.0	82.5	47.6	75.2	86.3	90.2
SSG [19]	ICCV 2019	53.4	73.0	80.6	83.2	58.3	80.0	90.0	92.4
AD-Cluster [13]	CVPR 2020	54.1	72.6	82.5	85.5	68.3	86.7	94.4	96.5
DG-Net++ [36]	ECCV 2020	63.8	78.9	87.8	90.4	61.7	82.1	90.2	92.7
MMT [21]	ICLR 2020	65.1	78.0	88.8	92.5	71.2	87.7	94.9	96.9
MEB-Net [32]	ECCV 2020	66.1	79.6	88.3	92.2	76.0	89.9	96.0	97.5
Dual-Refinement [49]	arXiv 2020	67.7	82.1	90.1	92.5	78.0	90.9	96.4	97.7
SSKD [50]	arXiv 2020	67.2	80.2	90.6	93.3	78.7	91.7	97.2	98.2
ABMT [51]	WACV 2021	70.8	83.3	–	–	80.4	93.0	-	-
RDSBN [52]	CVPR 2021	66.6	80.3	89.1	92.6	81.5	92.9	97.6	98.4
SECRET [53]	AAAI 2022	67.1	80.3	–	–	79.8	92.3	–	–
CLM-Net [54]	NCA 2022	69.7	82.3	90.5	93.2	80.9	92.4	97.3	98.3
LF2 [20]	ICPR 2022	73.5	83.7	91.9	94.3	83.2	92.8	**97.8**	98.4
HDNet [55]	IJMLC 2023	68.7	81.2	90.9	93.3	79.5	92.0	97.2	98.3
UMDA [56]	VCIR 2024	67.5	80.6	90.3	93.2	81.7	93.4	97.6	98.3
CORE-ReID (w/o ECAB)	Ours	74.3	84.7	**92.5**	94.2	83.2	93.6	97.7	98.5
CORE-ReID (w ECAB)	Ours	**74.8**	**84.8**	92.4	**94.4**	**84.4**	**93.6**	97.7	**98.7**

Bold denotes the best while Underline indicates the second-best results.

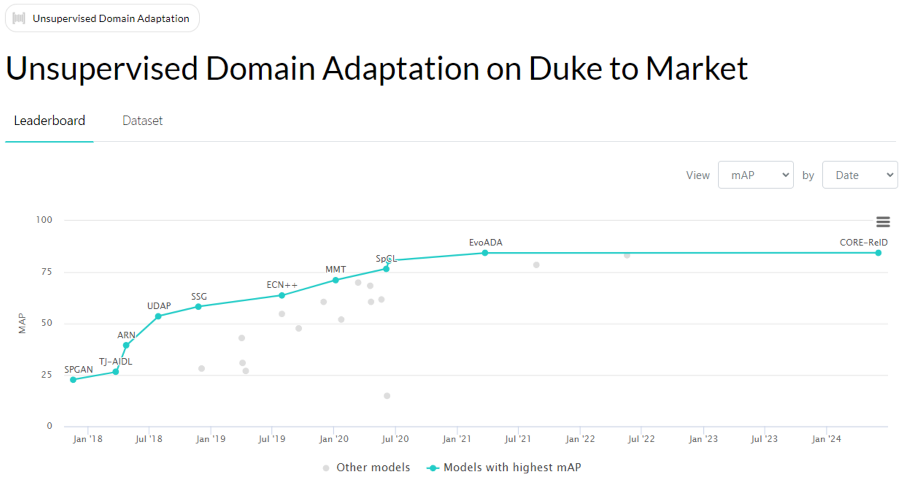
Ablation Study:: Additional Experiments

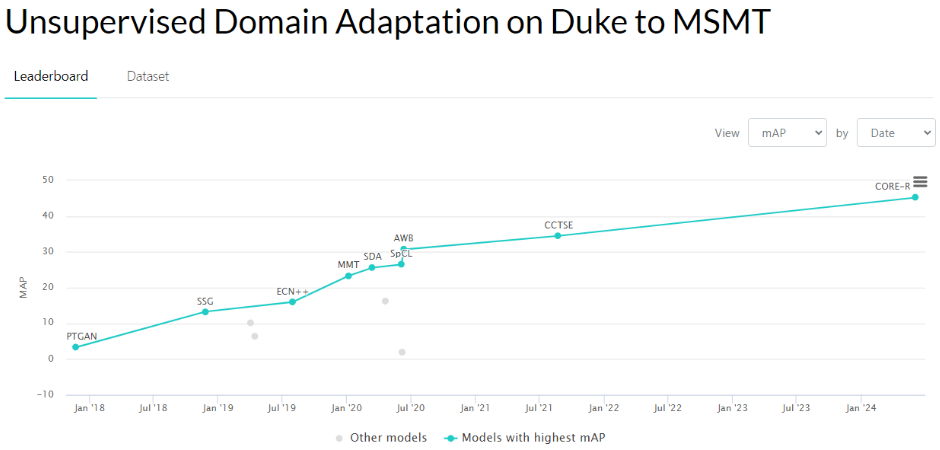
		Market To MSMT				Duke To MSMT			
Method	**Reference**	**mAP**	**R-1**	**R-5**	**R-10**	**mAP**	**R-1**	**R-5**	**R-10**
SSG [19]	ICCV 2019	13.2	31.6	-	49.6	13.3	32.2	-	51.2
MMCL [57]	CVPR 2020	15.1	40.8	51.8	56.7	16.2	43.6	54.3	58.9
NRMT [58]	ECCV 2020	19.8	43.7	56.5	62.2	20.6	45.2	57.8	63.3
DG-Net++ [36]	ECCV 2020	22.1	48.4	-	-	22.1	48.8	-	-
MMT [21]	ICLR 2020	22.9	52.5	-	-	22.9	50.1	-	-
Dual-Refinement [49]	arXiv 2020	25.1	53.3	66.1	71.5	26.9	55.0	68.4	73.2
RDSBN [52]	CVPR 2021	30.9	61.2	73.1	77.4	33.6	64.0	75.6	79.6
CLM-Net [54]	NCA 2022	29.0	56.6	69.0	74.3	26.6	53.8	65.2	70.7
HDNet [55]	IJMLC 2023	25.9	53.4	66.4	72.1	26.8	54.6	70.9	73.0
DDNet [59]	AI 2023	28.5	59.3	72.1	76.8	31.4	63.8	75.1	79.3
CaCL [60]	ICCV 2023	36.5	66.6	75.3	80.1	-	-	-	-
OUDA [61]	WACV 2024	20.2	46.1	-	-	-	-	-	-
M-BDA [62]	VCIR 2024	26.7	51.4	64.3	68.7	23.4	47.3	59.5	64.5
UMDA [56]	VCIR 2024	32.7	62.4	72.7	78.4	34.1	64.7	76.2	80.5
CORE-ReID (w/o ECAB)	Ours	<u>41.2</u>	<u>68.5</u>	<u>80.1</u>	<u>83.8</u>	<u>44.6</u>	<u>72.2</u>	<u>82.8</u>	<u>86.2</u>
CORE-ReID (w ECAB)	Ours	**41.9**	**69.5**	**80.3**	**84.4**	**45.2**	**72.2**	**82.9**	**86.3**

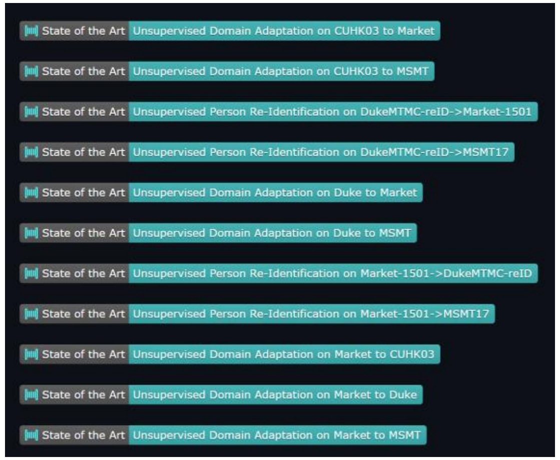
Bold denotes the best while <u>Underline</u> indicates the second-best results.

Iwate Prefectural University

Ablation Study:: Benchmark on PaperWithCode







#	Source dataset	Target dataset	Paper with code (CORE-ReID)	Rank	Note
1	DukeMTMC	Market-1501	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-duke-to	Top1	
2	DukeMTMC	MSMT17	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-duke-to-1	Top1	
3	CUHK03	Market-1501	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-cuhk03-to-1	Top1	New
4	CUHK03	MSMT17	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-cuhk03-to	Top1	New
5	Market-1501	CUHK03	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-market-to-6	Top1	New
6	Market-1501	DukeMTMC	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-market-to	Top1	
7	Market-1501	MSMT17	https://paperswithcode.com/sota/unsupervised-domain-adaptation-on-market-to-1	Top1	

Conclusion

Achievements and contributions

- Proposed a dynamic fine-tuning strategy using a camera-aware style transfer model to reduce camera style disparities and prevent CNN overfitting.
- Introduced Efficient Channel Attention Block (ECAB) to enhance feature extraction by prioritizing meaningful structures.
- Developed the CORE-ReID framework, which employs teacher-student networks and Ensemble Fusion component to fuse global and local features to improved pseudo-label generation.
- Incorporated Bidirectional Mean Feature Normalization (BMFN) to improve feature discriminability.
- Achieved state-of-the-art (SOTA) performance and reduced the gap between supervised and unsupervised Person Re-ID.

Future Work

Limitations and Solutions

1 Only local features are enhanced using ECAB in the Ensemble Fusion

Improve the Ensemble Fusion component by enhancing global feature as well

2 CORE-ReID only consider Person ReID

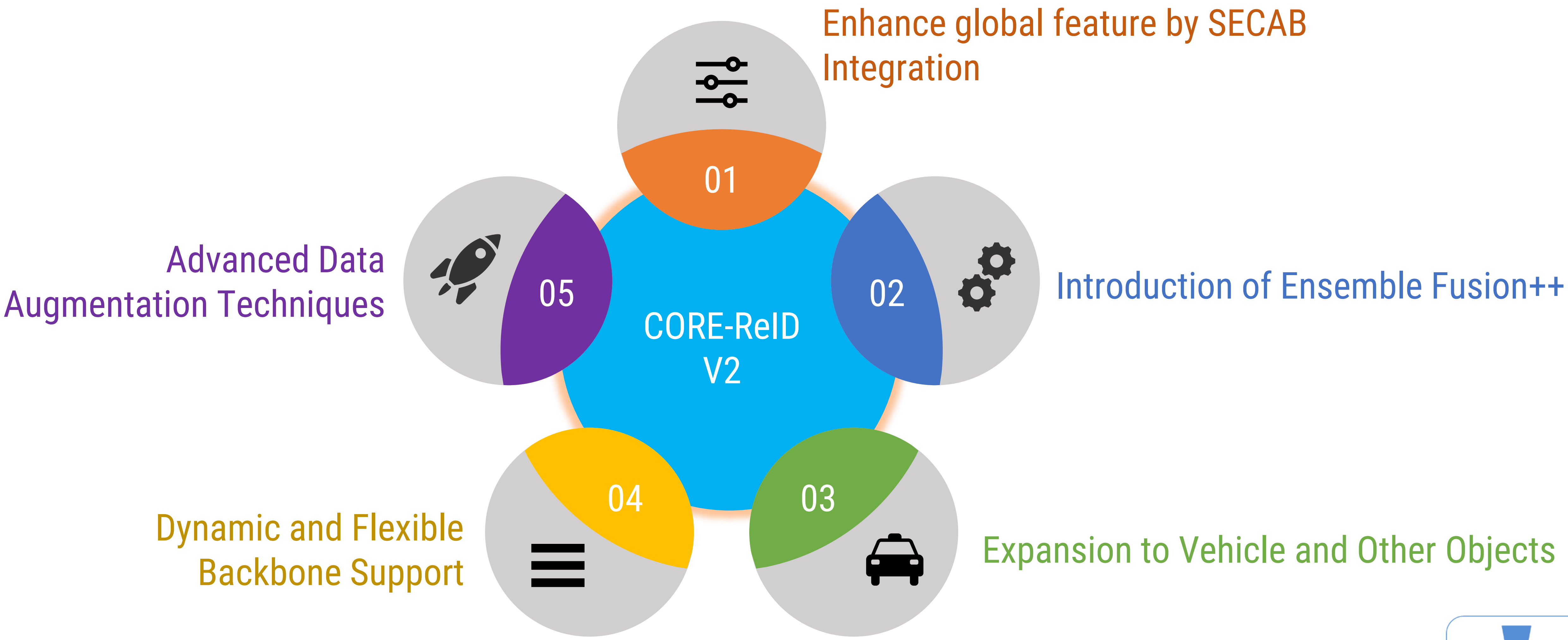
Expand to Vehicle ReID, and further Object ReID

3 Only ResNet 50, 101, 152 are currently supported

Support small backbones

CORE-ReID V2

Future Work



Enhance global feature by SECAB Integration
01
Introduction of Ensemble Fusion++
02
Expansion to Vehicle and Other Objects
03
Dynamic and Flexible Backbone Support
04
Advanced Data Augmentation Techniques
05
CORE-ReID V2

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] Ni, X.; Rahtu, E. FlipReID: Closing the Gap between Training and Inference in Person Re-Identification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021.

[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.

[19] Wei, L.; Zhang, S.; Gao, W.; Tian, Q. Person transfer gan to bridge domain gap for person reidentification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 79–88.

[20] Nguyen, T.Q.; Prima, O.D.A.; Hotta, K. CORE-ReID: Comprehensive Optimization and Refinement through Ensemble Fusion in Domain Adaptation for Person Re-Identification. Software 2024, 3, 227-249. https://doi.org/10.3390/software3020012.

Thank you