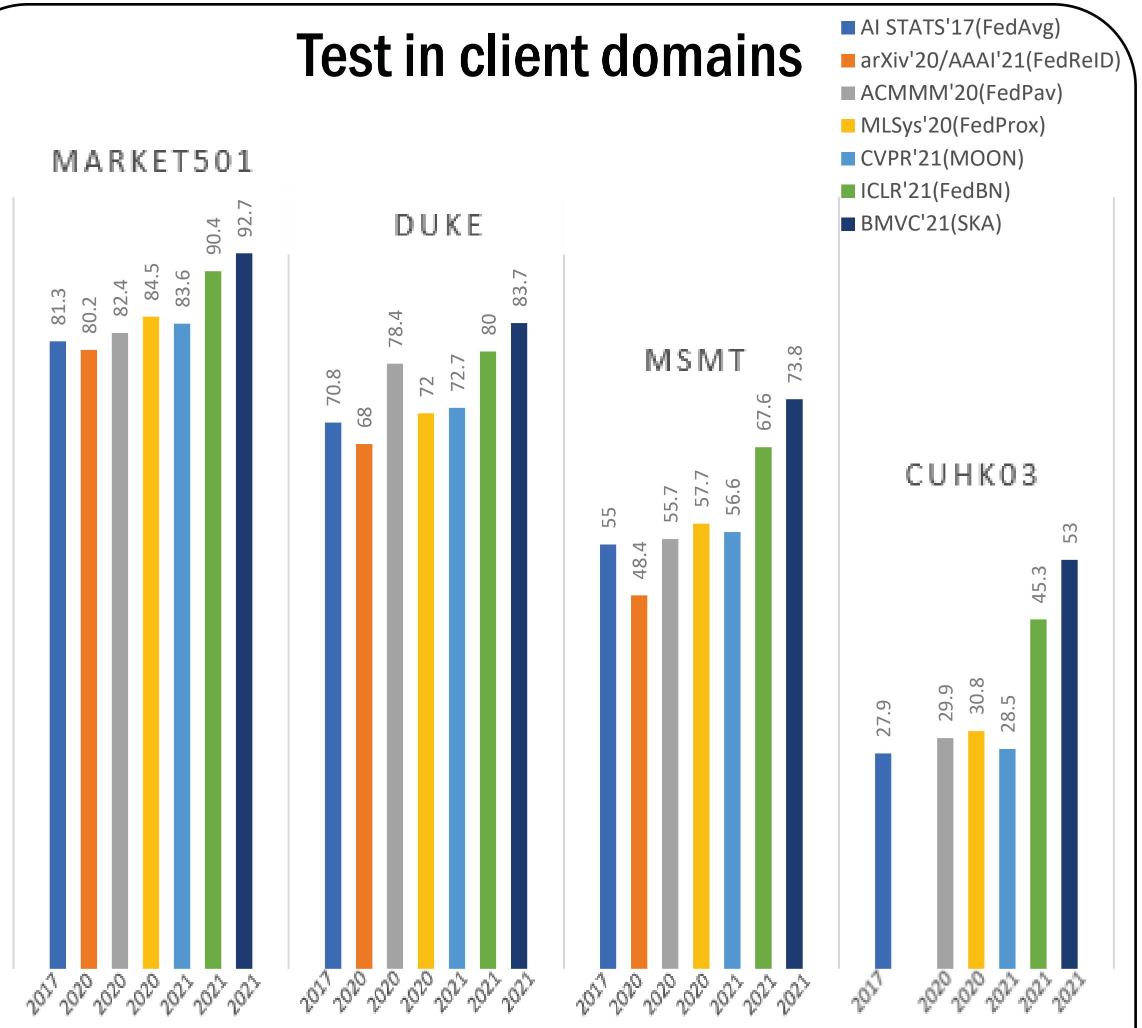
Decentralised Person Re-Identification with Selective Knowledge Aggregation

Shitong Sun, Guile Wu, and Shaogang Gong

Queen Mary University of London

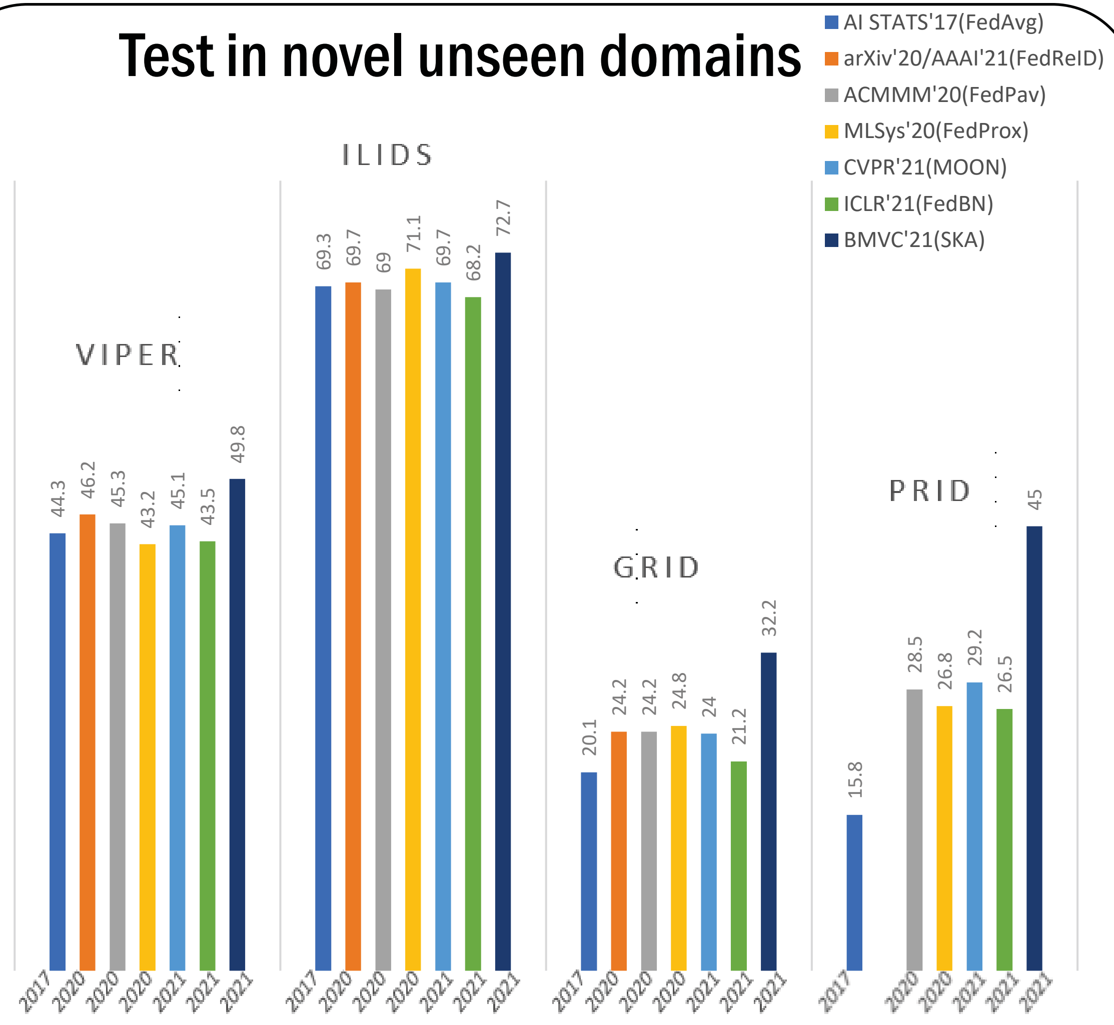
The Success of Federated Person Re-ID (reproduced result)

Test in client domains



Improved on average by +29% (2017-2021)

Test in novel unseen domains



Improved on average by +33%(2017-2021)

Limitations of Federated Learning

Personalisation → Collecting multi-domain datasets for decentralised training & applied to seen source domains

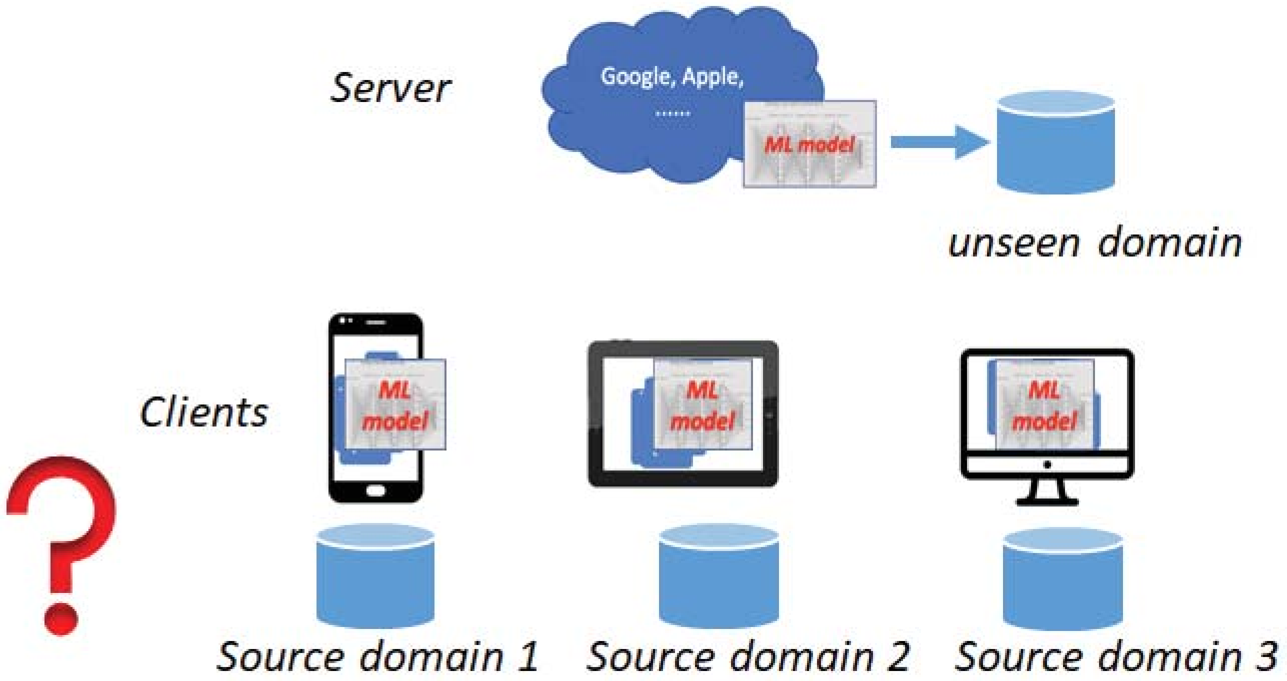
Generalisation → Assembling multi-domain datasets for decentralised training, then applied to unseen target domains

Challenges:

How to train a model when

- specific knowledge extraction for each individual domain/client
- general knowledge generalisation for unseen domain

are both necessary due to real world requirement?



Related Works

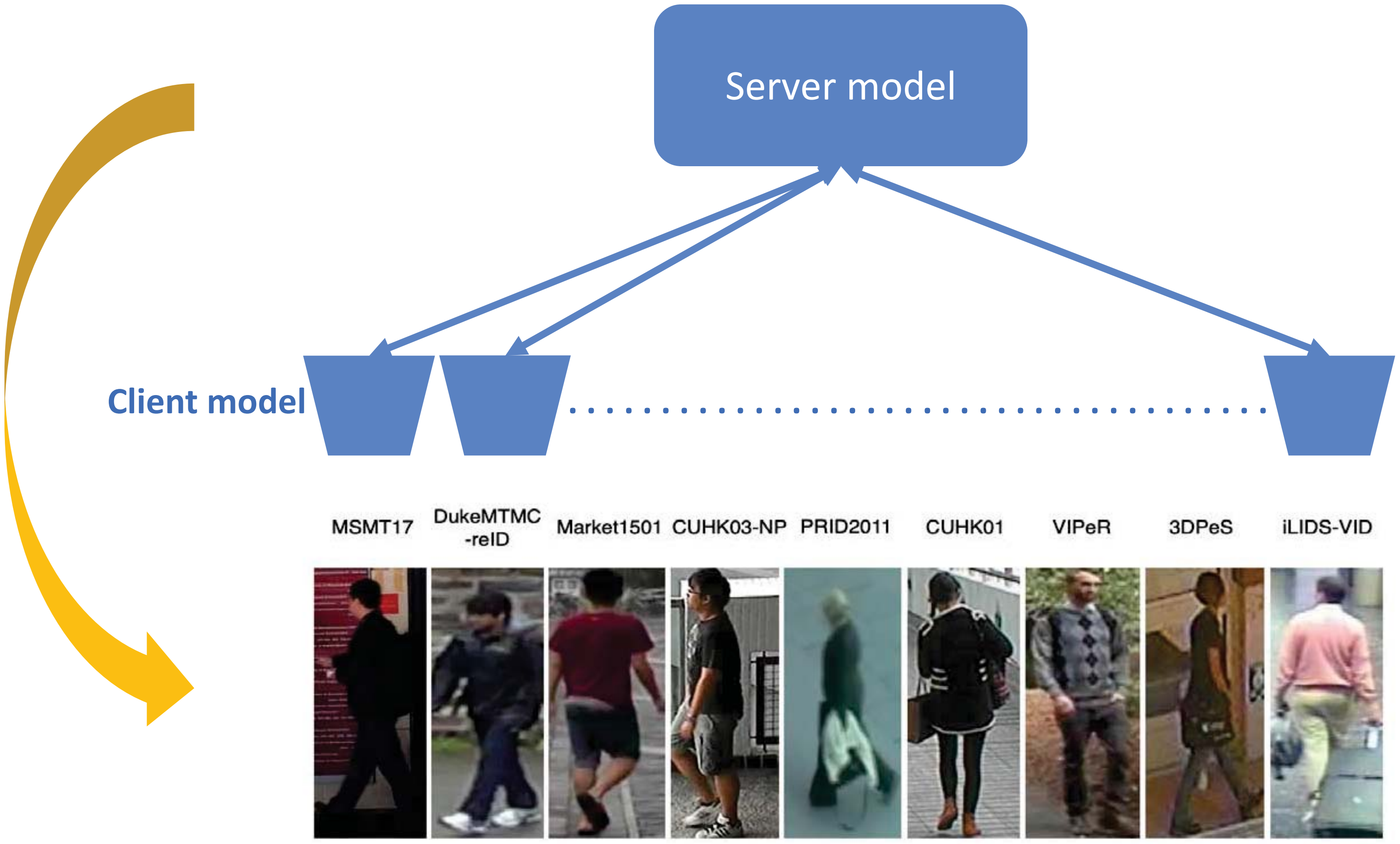
Federated Learning (Konecny et al. arXiv 2016) – Enables localised users to collaboratively train a centralised model without sharing local data, but assuming a shared common label space (e.g. CIFAR) & with labelled training data (N-Shot Learning, not ZSL)

J Konecny, H B McMahan, F X Yu, P Richtarik, A T Suresh, D Bacon, Federated learning: Strategies for improving communication efficiency, arXiv:1610.05492, 2016.

Federated Person ReID(e.g. Wu et al AAAI 2021, Zhuang et al. ACM 2020, Zhuang ACM 2021) - relies on decentralised labelled/unlabelled training data, assuming generalisable global/personalised local model can be learned from decentralised data with privacy protection.

Guile Wu, Shaogang Gong, Decentralised Learning from Independent Multi-domain Labels for Person Re-Identification, AAAI 2021.
Weiming Zhuang et al, Performance Optimization for Federated Person Re-identification via Benchmark Analysis, ACMMM 2020.
Weiming Zhuang et al, Joint Optimization in Edge-Cloud Continuum for Federated Unsupervised Person Re-identification, ACMMM 2021.

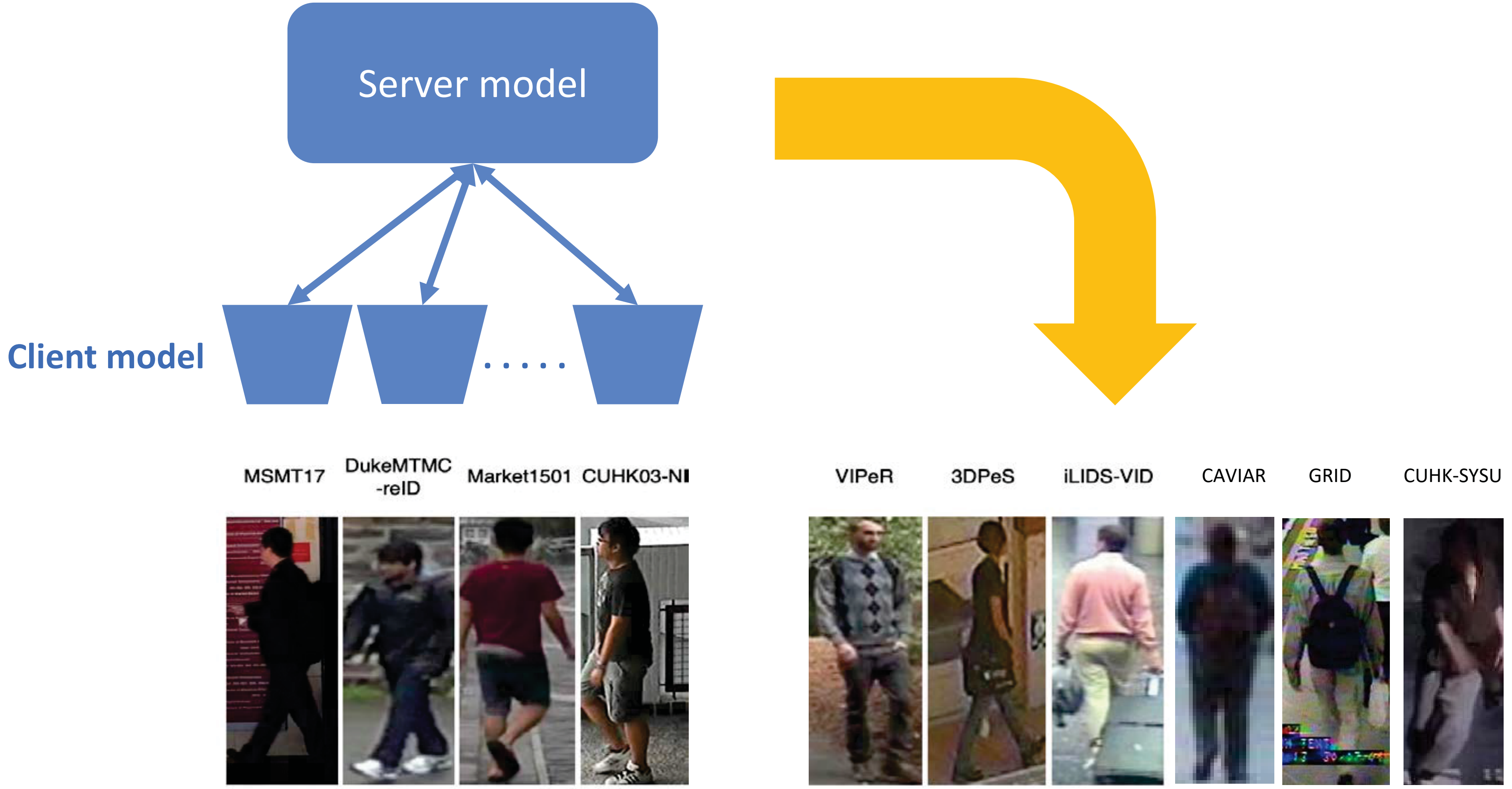
FedPav



Only Personalisation

Weiming Zhuang et al, Performance Optimization for Federated Person Re-identification via Benchmark Analysis, ACMMM 2020.

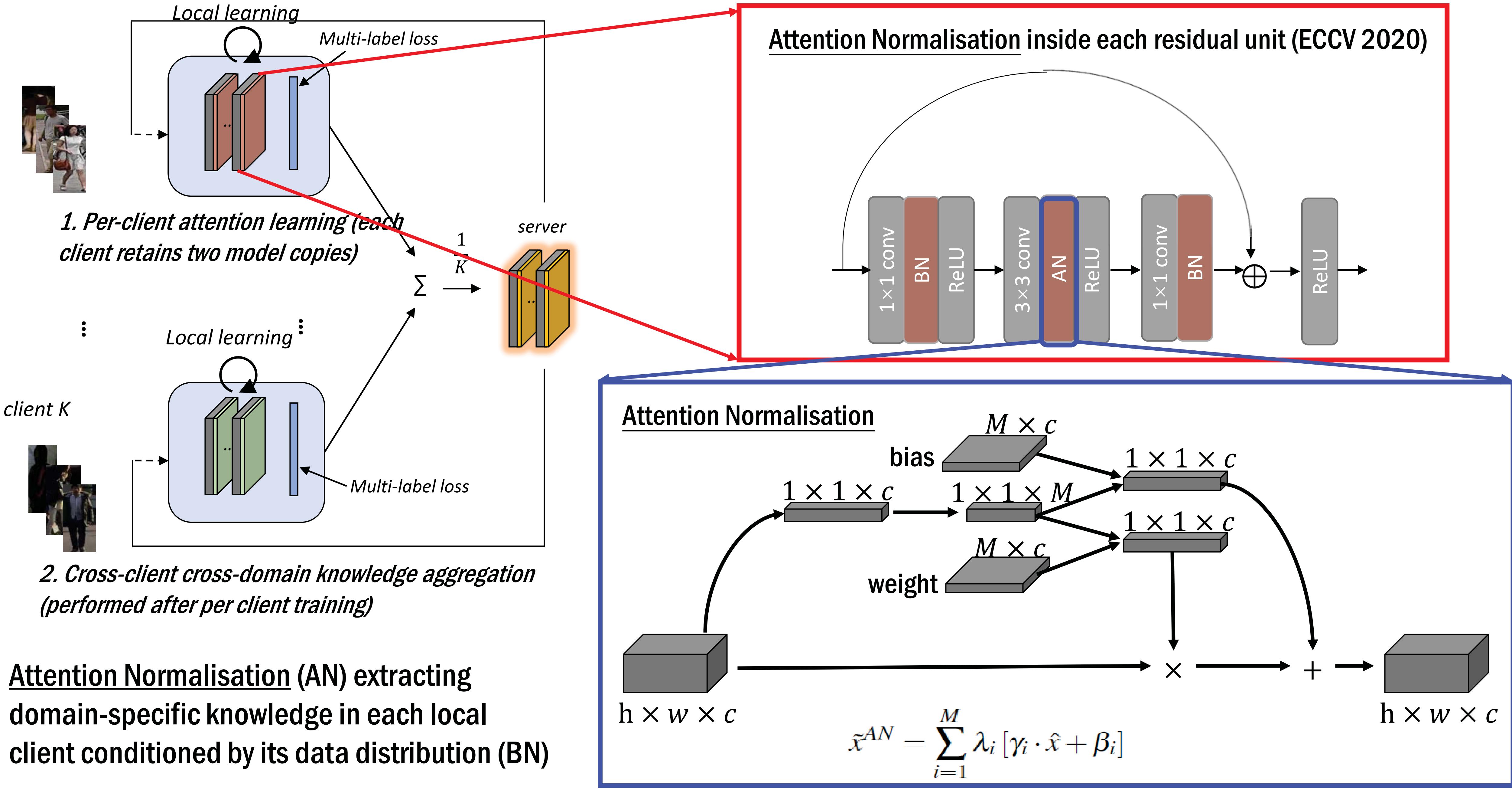
FedReID



Only Generalisation

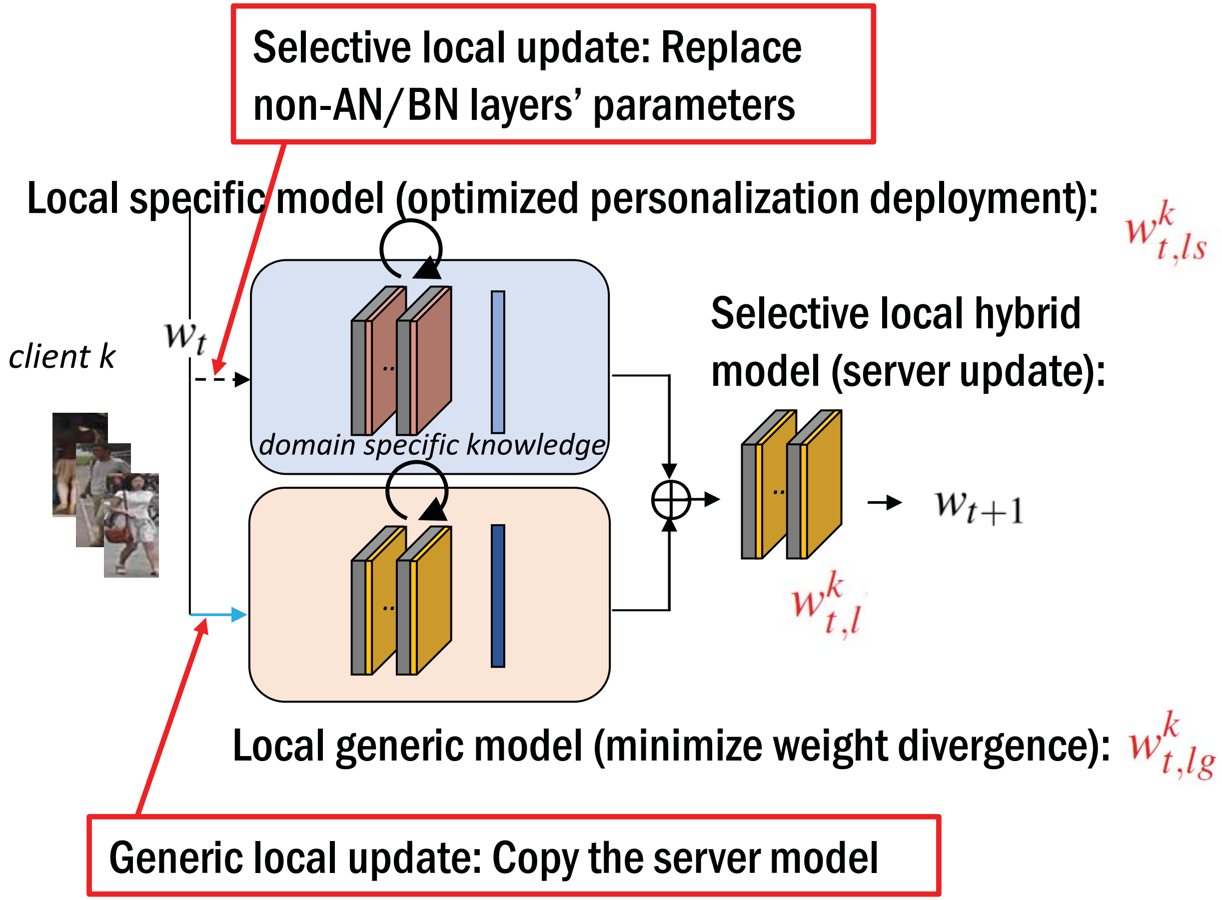
Guile Wu, Shaogang Gong, Decentralised Learning from Independent Multi-domain Labels for Person Re-Identification, AAAI 2021.

Key idea 1: Per-client specific attention normalisation



Attention Normalisation (AN) extracting domain-specific knowledge in each local client conditioned by its data distribution (BN)

Key idea 2: Cross-client knowledge aggregation across data distributions



Asynchronous global training cycle:
Selective local update & selective global hybrid per client after each client local training: Local epoch (1) ×mini batch number (avoid weight divergence)

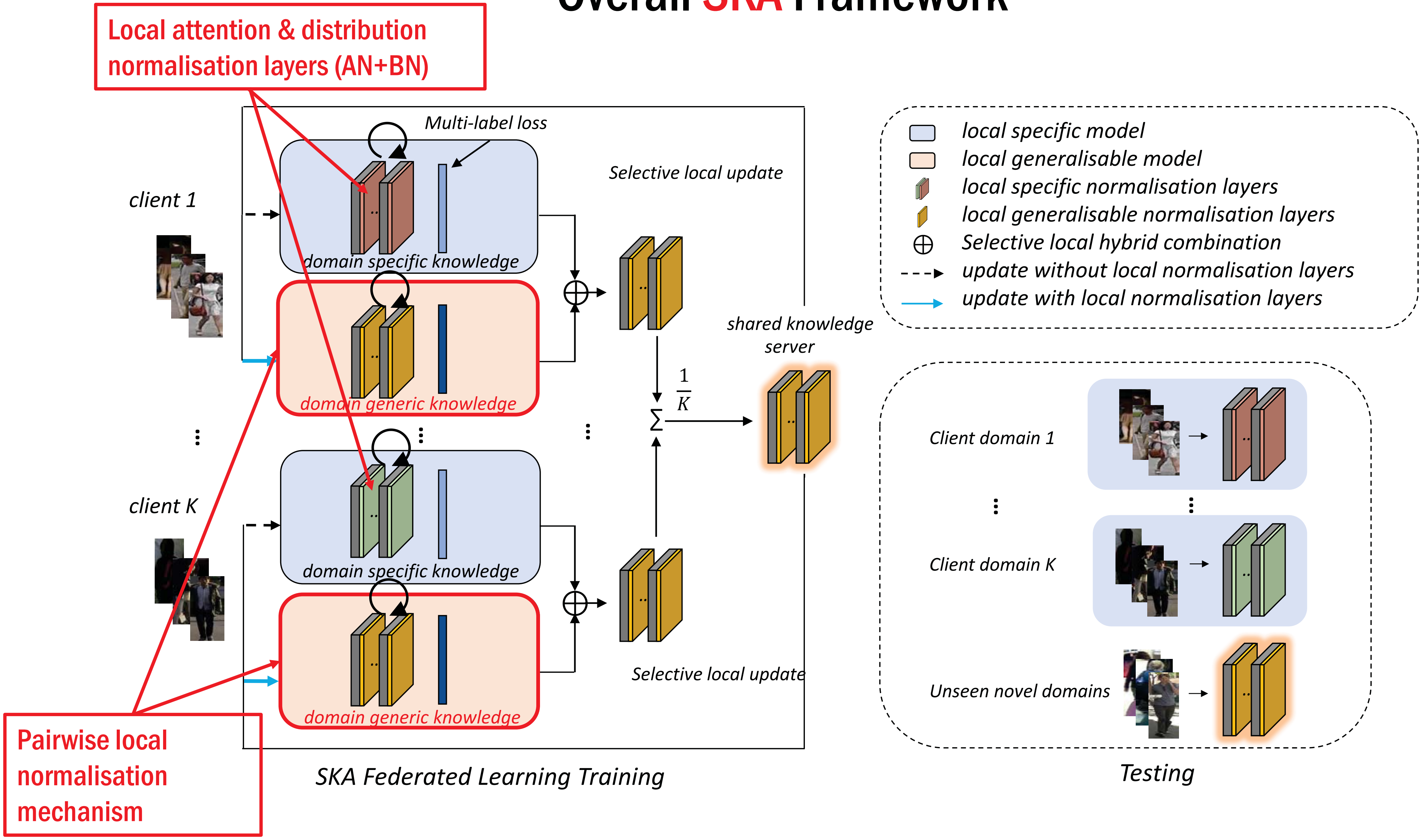
Selective local update:

Update without normalisation layers, keep local knowledge within each client for optimised local model

Selective local hybrid:

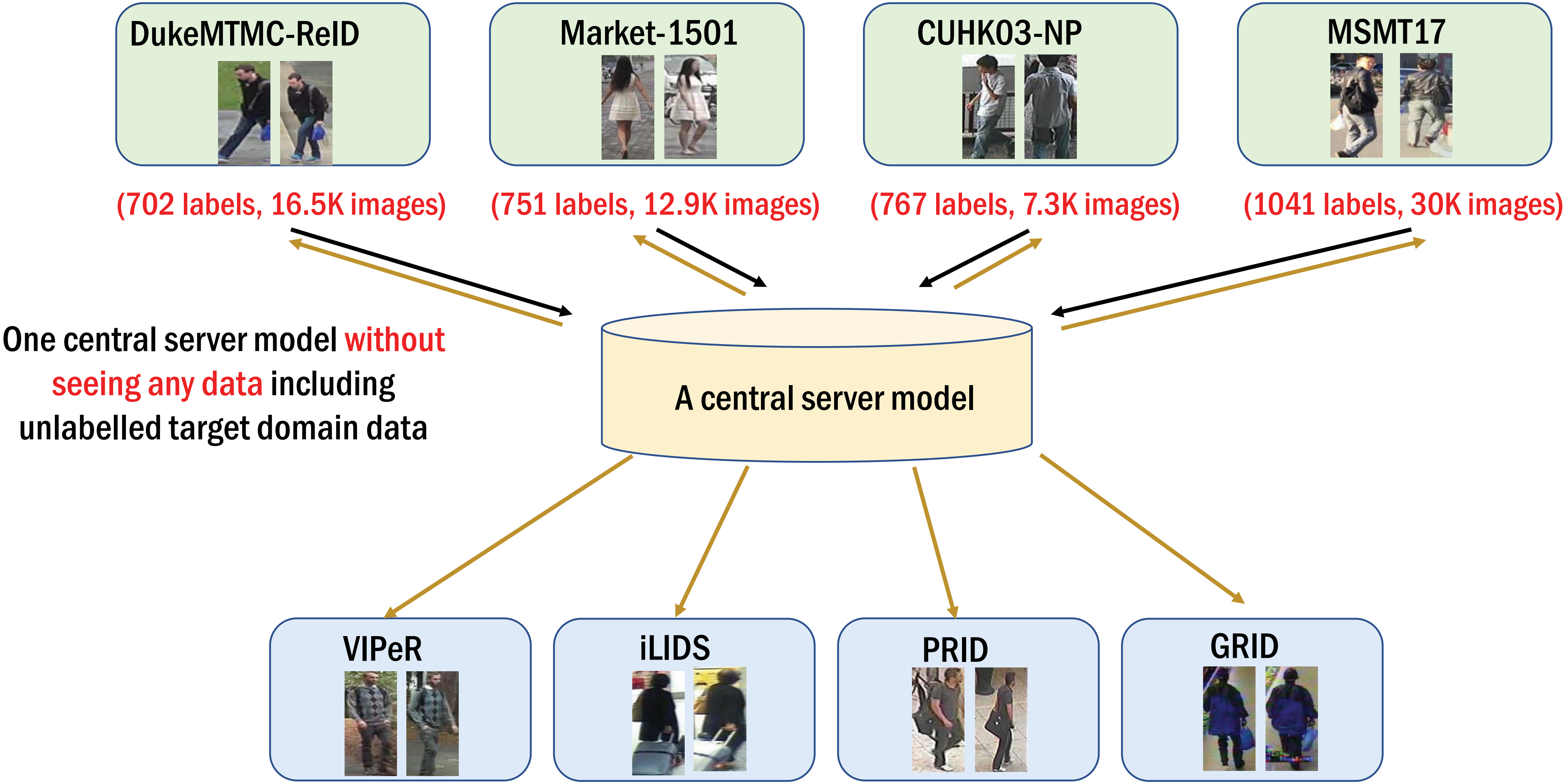
Selective combination between non-normalisation layers of the local specific model and all other layers from the local generic model

Overall SKA Framework



Test

Four independent local client models without sharing training data



Four unseen target domains for 'out-of-the-box' tests (decentralised client training domains are NOT the test domains)

Test

Method	Source	Duke		Market1501		CUHK03-NP		MSMT17	
		mAP	R1	mAP	. R1	mAP	R1	mAP	R1
local supervised		57.0	77.2	68.2	87.4	39.2	43.2	28.8	60.0
FedAvg[21]+D	AISTATS'17	50.9	70.8	56.6	81.3	25.2	27.9	25.9	55.0
FedProx[13]+D	MLSys'20	53.5	72.0	59.7	84.5	27.8	30.8	28.1	57.7
FedPav[41]	ACMMM'20	51.9	71.1	53.5	78.7	23.0	26.0	26.1	54.4
FedPav+AddData[41]	ACMMM'20	60.6	78.4	58.0	82.4	26.8	29.9	27.0	55.7
FedReID[32]	AAAI'21	52.1	68.0	60.1	80.2	-	-	-	48.4
MOON [12]+D	CVPR'21	53.1	72.7	58.1	83.6	26.4	28.5	27.2	56.6
FedBN[16]+D	ICLR'21	62.8	80.0	73.1	90.4	40.9	45.3	35.4	67.6
SKA	Ours	**66.6**	**83.7**	**78.2**	**92.7**	**48.1**	**53.0**	**42.9**	**73.8**

Table 2: Comparisons of decentralised learning Re-ID on four source client domains.

Method	Source	VIPeR		iLIDS		GRID		PRID	
		mAP	R1	mAP	R1	mAP	R1	mAP	R1
FedAvg[21]+D	AISTATS'17	48.2	44.3	73.3	69.3	24.3	20.1	19.6	15.8
FedProx[13]+D	MLSys'20	47.3	43.2	74.9	71.1	29.1	24.8	31.2	26.8
FedPav[41]	ACMMM'20	49.5	44.9	72.8	68.8	25.5	21.7	37.0	31.9
FedPav+AddData[41]	ACMMM'20	49.6	45.3	73.1	69.0	28.7	24.2	34.4	28.5
FedReID[32]	AAAI'21	-	46.2	-	69.7	-	24.2	-	-
MOON [12]+D	CVPR'21	49.1	45.1	73.7	69.7	28.0	24.0	33.5	29.2
FedBN[16]+D	ICLR'21	47.9	43.5	72.3	68.2	25.2	21.2	31.1	26.5
SKA	Ours	**53.9**	**49.8**	**76.0**	**72.7**	**36.7**	**32.2**	**49.7**	**45.0**

Table 3: Comparisons of decentralised learning Re-ID on four unseen novel domains.

Test

Model personalisation improves by AN and local specific normalisation

Model generalisation improves on unseen datasets by dual mechanism

Components	Seen domains				Unseen domains			
	Duke		Market		VIPeR		iLIDS	
	mAP	R1	mAP	R1	mAP	R1	mAP	R1
Baseline	50.9	70.8	56.6	81.3	48.2	44.3	73.3	69.3
Baseline+AN	58.7	78.1	64.2	87.9	54.2	50.2	77.0	74.0
Baseline+LSN	62.8†	80.0†	73.1†	90.4†	47.9	43.5	72.3	68.2
Baseline+LSN+Dual	62.8†	80.0†	73.1†	90.4†	48.8	44.6	75.2	71.3
Baseline+LSN+AN	66.6*	83.7*	78.2*	92.7*	51.8	47.8	73.1	69.3
Baseline+LSN+AN+Dual	66.6*	83.7*	78.2*	92.7*	53.9	49.8	76.0	72.7

Table 4: Evaluating component effectiveness on seen source domains and unseen novel domains. *,† 'Dual' does not affect model personalisation on seen source domains as local generalised normalisation layers are only learnt to construct the global generalised model.

Variants	Duke	Market
SKA	83.7	92.7
SKA w/o LSN	78.1	87.9
SKA w/o AN	80.0	90.4
SKA w/o AN + SE	82.0	91.2
SKA w/o AN + CBAM	80.5	91.1

Table 5: Evaluating variants of learning local personalisation knowledge (R1).

Variants	VIPeR	iLIDS
SKA w/ Dual	49.8	72.7
SKA w/o Dual + Avg	47.8	69.3
SKA w/o Dual + FeatConcat	47.2	71.8
SKA w/o Dual + RandImg	6.6	21.8

Table 6: Evaluating variants of dual local normalisation (R1).

Takeaway

"*Decentralised AI for augmenting human intellect at the edge*"

Current Federated Learning Paradigm

- Distributed model learning with pricavy protection
- Only focusing on personalisation on local clients
- Only focusing on generalisation on global server

Our Selective Knowledge Aggregation:

- Balance the trade-off between personalisation and generalisation
- Normalisation based method allows knowledge selection

Thanks