Deep Unsupervised Progressive Learning for Distant Domain Adaptation

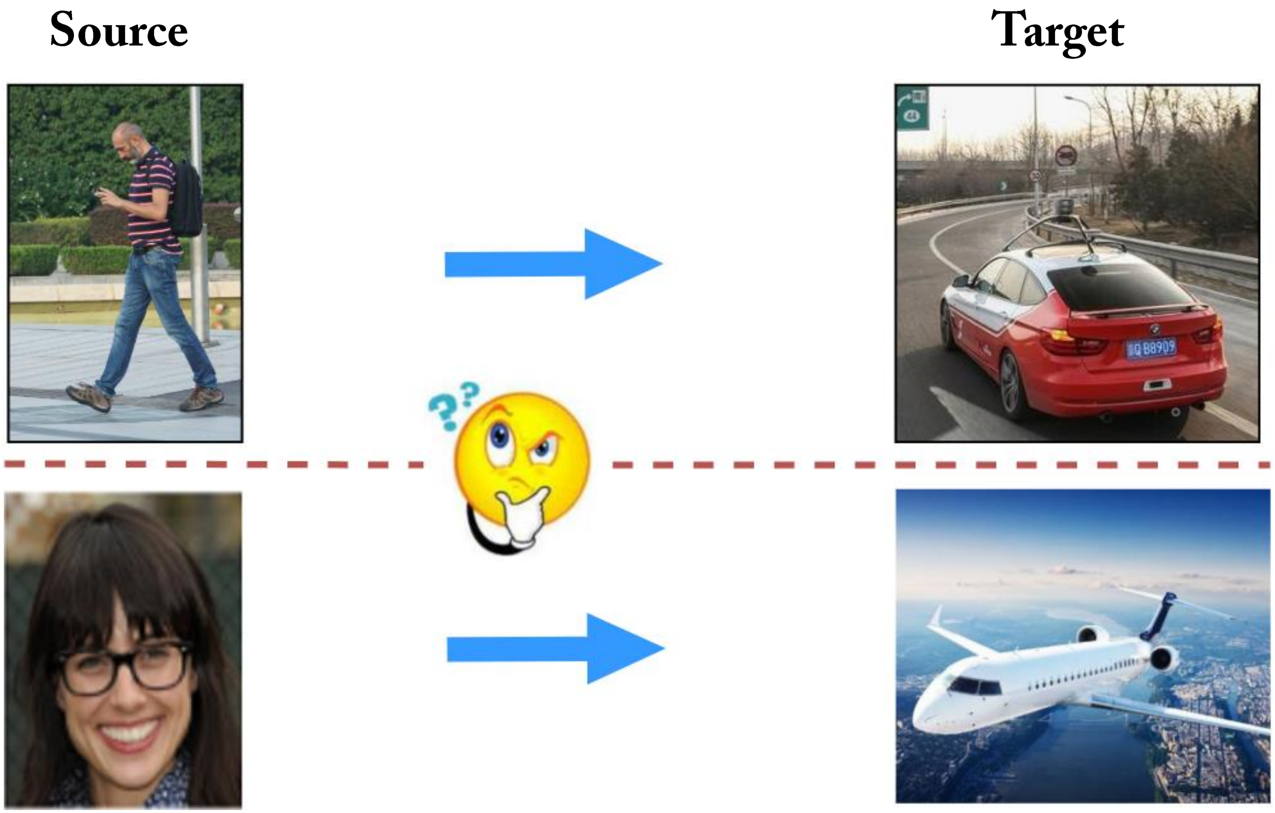
Suncheng Xiang, Yuzhuo Fu, Ting Liu

School of Microelectronics
Shanghai Jiao Tong University

Introduction: Distant Domain Adaptation

Distant Domain Adaptation: Transfer between two distant domain

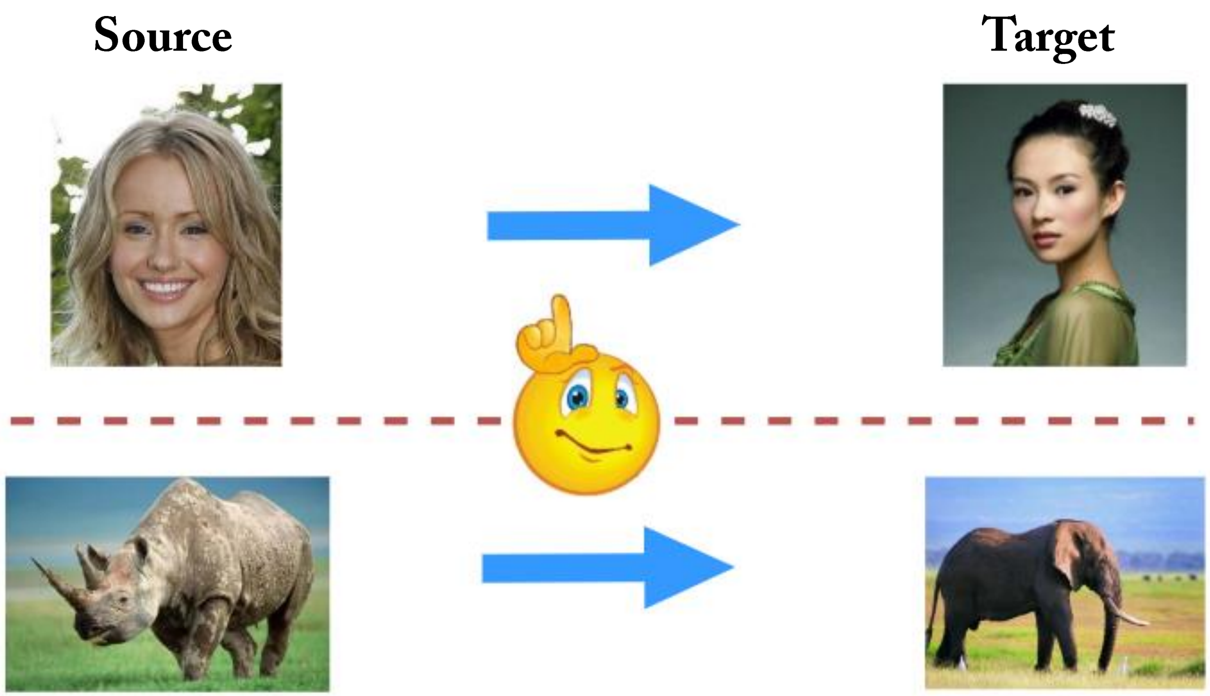
e.g. Classify **vehicle images** with a model trained on **pedestrian images**



Related Work

Domain Adaptation: Transfer between two close domain (requirement)

e.g. Classify **human faces** or **animals** with similar appearance



Related Work

Domain Adaptation:

Source domain **~ $P_S(X,Y)$** $\neq$ ***Target domain*** **~ $P_T(Z,H)$**

lots of **labeled** data

unlabeled or limited labels

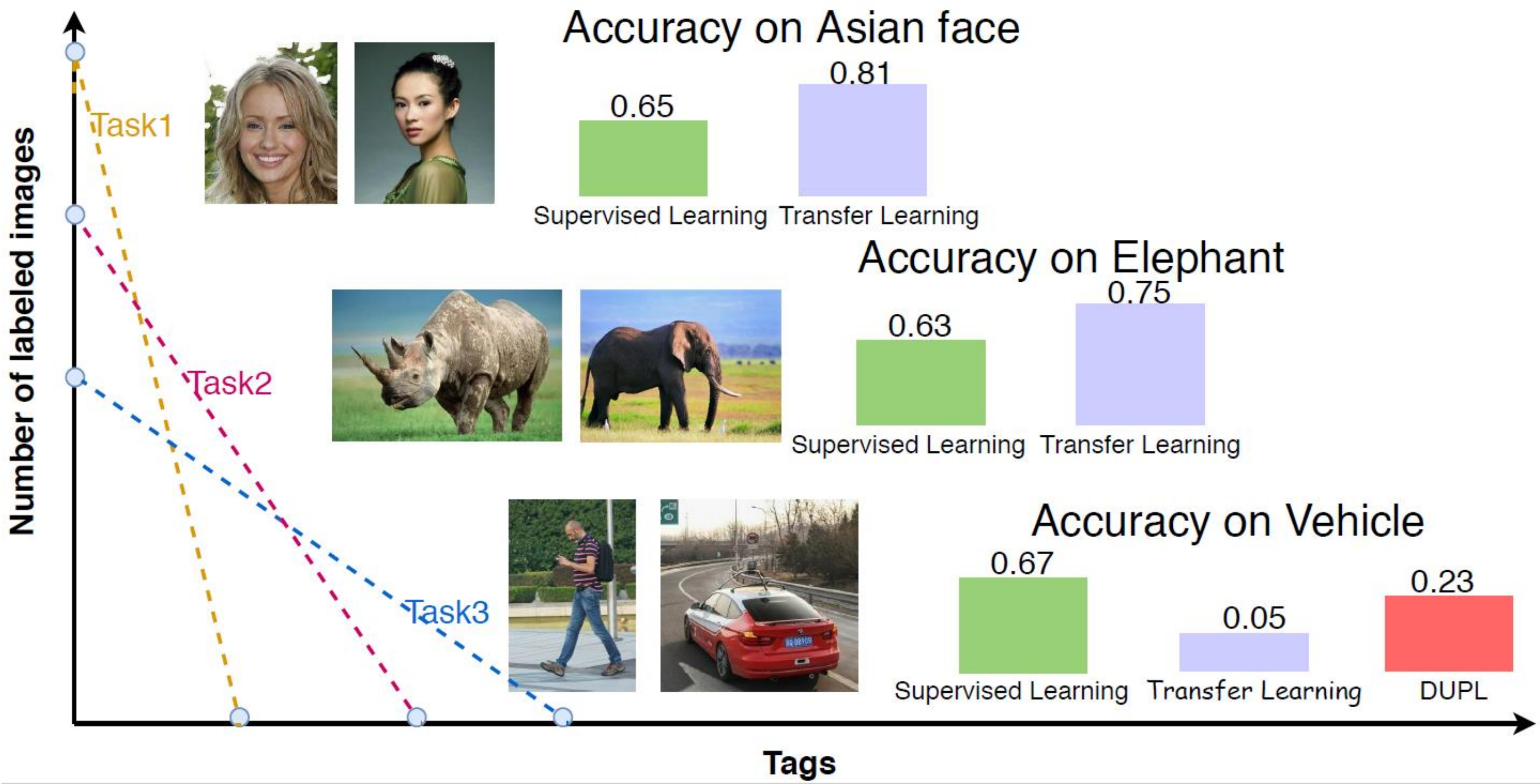
D_S={($\boldsymbol{x}_i$, y_i), $\forall i \in$ {1,…,N}}

D_T={($\boldsymbol{z}_i$,**?**), $\forall j \in$ {1,…,M}}

- Find a **mapping function** from source domain to target domain
- **Domain-invariant** representations can be shared between two domains

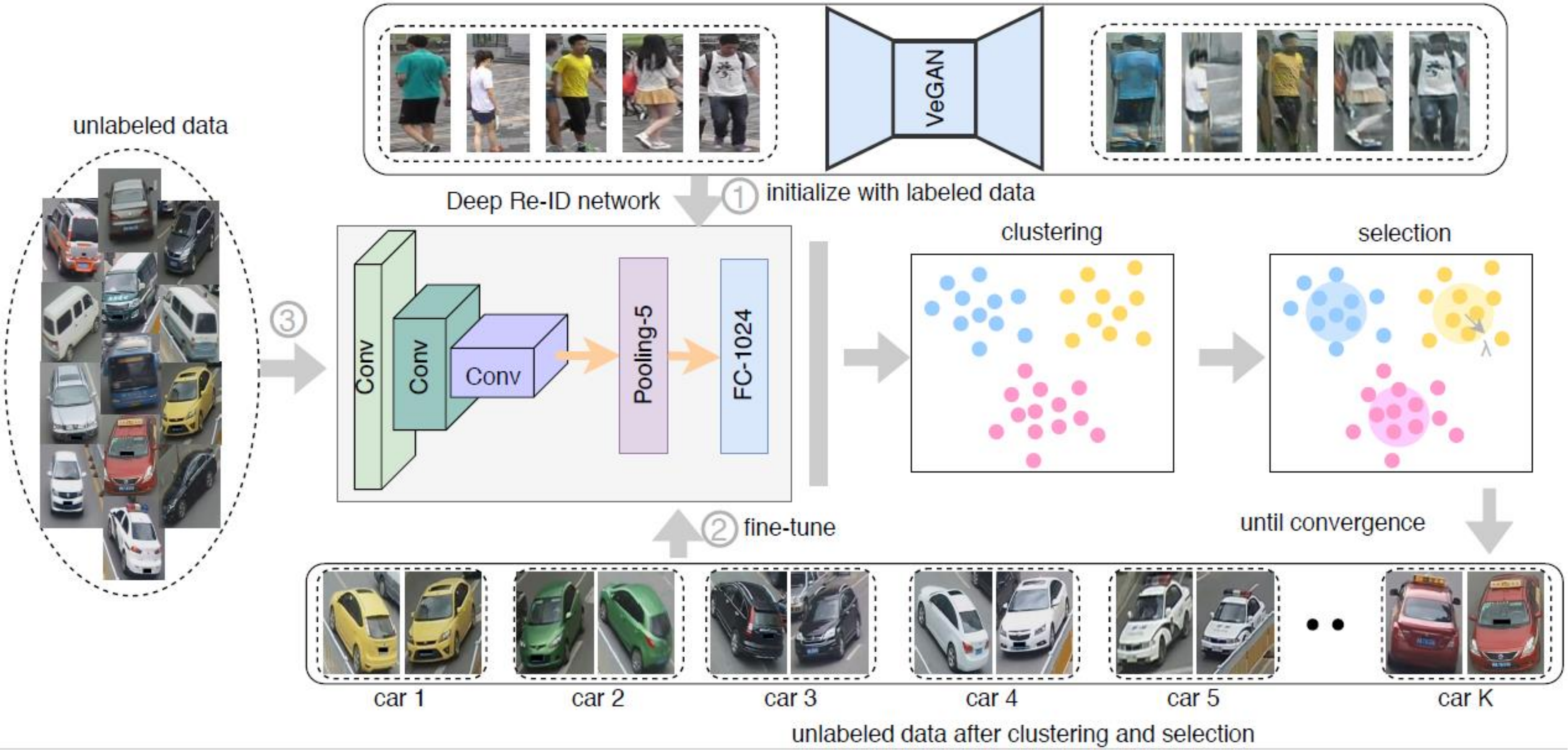
Our work

- ✓ Execute an unrelated task based on a labeled dataset training previously
- ☐ Challenging due to different distribution between source and target domain
- ☐ Worse performance than supervised learning



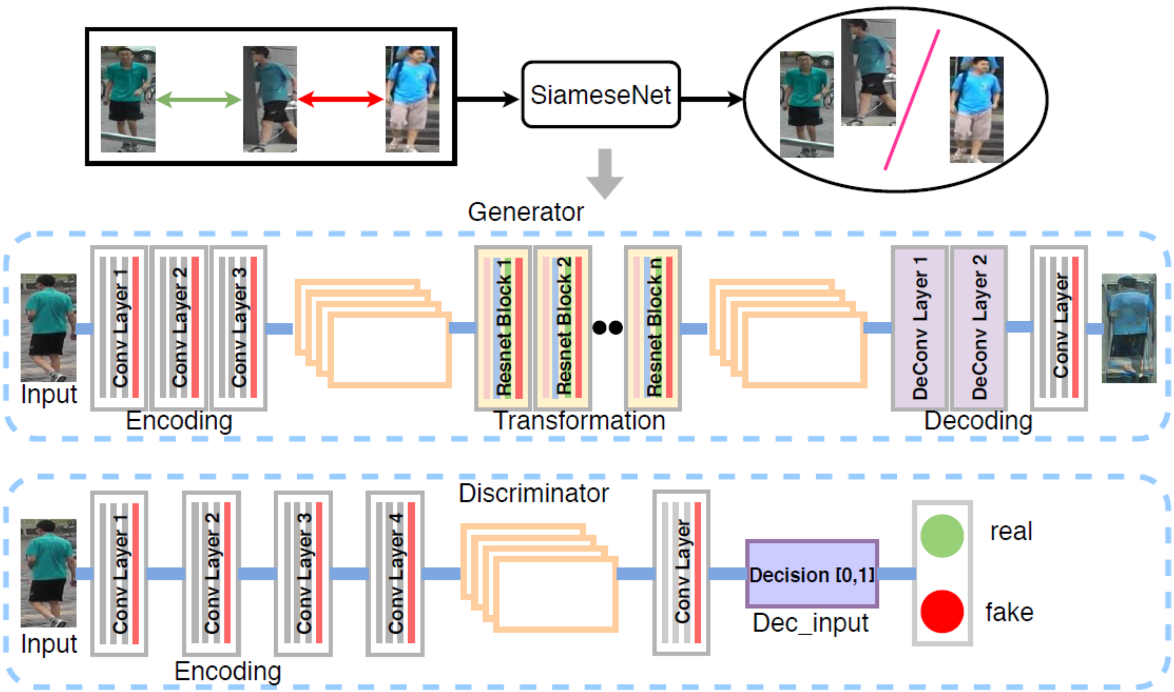
Solution: DUPL Framework

- Bridge the gap between source and target domain
- Mine the discriminative information (potential label information) in the target domain

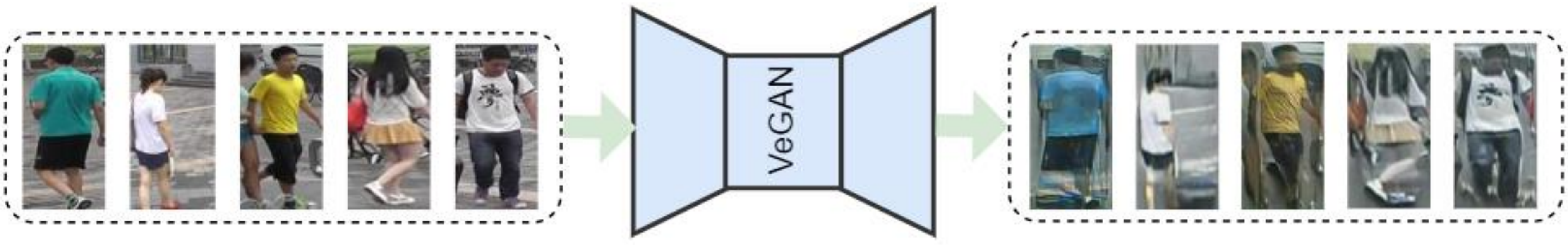


Solution: VeGAN Model

- Slightly modification based on the CycleGAN (image-image translation)
- Integrate a SiameseNet with CycleGAN (Similarity preserving learning)
- CycleGAN ➡ learning mapping function between two domain
- SiameseNet ➡ learning a latent space to constrain the mapping function



Solution: VeGAN Model



Objective function of VeGAN

$$\mathcal{L}_{VeGAN} = \mathcal{L}_{cyc} + \beta \mathcal{L}_{\text{ide}} + \gamma \mathcal{L}_{\text{con}}$$

- β and γ control the relative importance of the identity loss and constraint loss
- Training pairs are constructed in an unsupervised manner without sample annotation
- Training VeGAN until convergence or reaching maximum iterations

Solution: VeGAN Model

- Some visutal examples with our VeGAN
- Transferred sample has some latent information about target domain

Solution: Unsupervised Progressive Learning

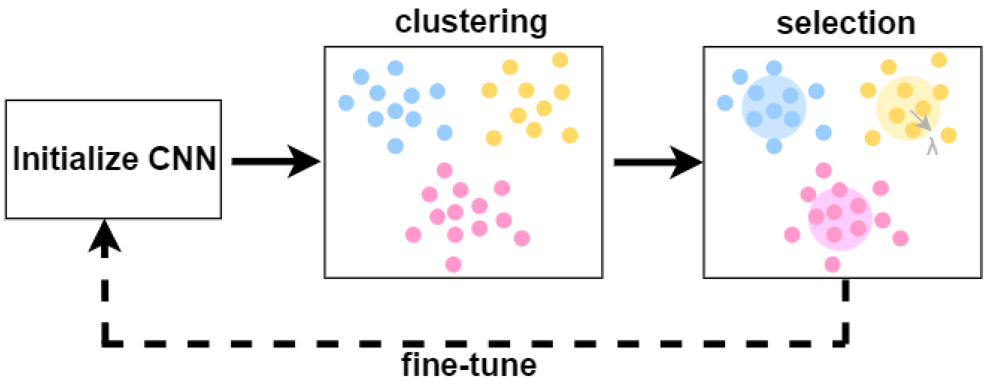
- Mine the potential label information in target domain

Algorithm 1 Unsupervised Progressive Learning

Input:
Irrelavant labeled data $\{x_i\}_{i=1}^{N}$;
No. of clusters K;
Base model $\phi(x, \theta_0)$

Output:
Fine tuned model $\phi(x, \theta_t)$

1: **initialize** $\theta_o \rightarrow \theta_t$
2: **while** not converged **do**
3: extract features $\mathbf{f}_i = \phi(x_i, \theta)$ for all $i = 1, \ldots, N$
4: cluster center $\mathbf{C} \leftarrow \arg\min \sum_{i=1}^{N} \sum_{j=1}^{M} \|\mathbf{f}_i - \mathbf{c}_j\|^2$
5: reliability constraint $\frac{\mathbf{f}_i}{\|\mathbf{f}_i\|} \bullet \frac{\mathbf{c}_j}{\|\mathbf{c}_j\|} > \lambda$
6: optimize $\min_{\theta,\mathbf{w}} \sum_{i=1}^{n} L((\mathbf{y}_i, \phi(\mathbf{x}_i, \theta)), \mathbf{w})$
7: **end while**
8: **return** $\rightarrow \theta_t$



Experiment Setting

- Training: Duke / Market / CUHK03
- Testing: VeRi dataset (ICME2016)

- VeGAN

$$\mathcal{L}_{VeGAN} = \mathcal{L}_{cyc} + \beta\mathcal{L}_{\mathrm{ide}} + \gamma\mathcal{L}_{\mathrm{con}}$$

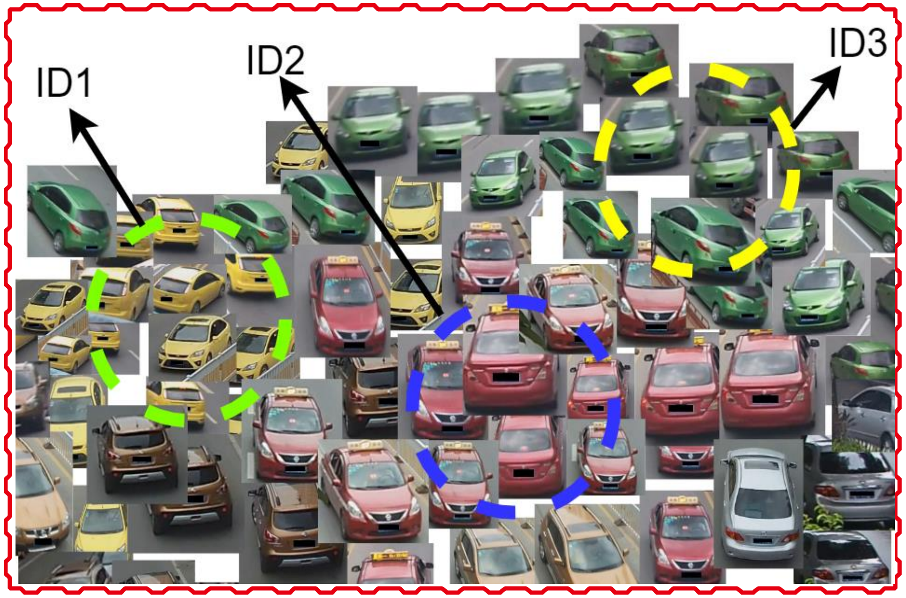
β = 5, γ = 2; learning rate = 0.0002 ; training 6 epochs

- DUPL learning framework

Modified from ResNet-50

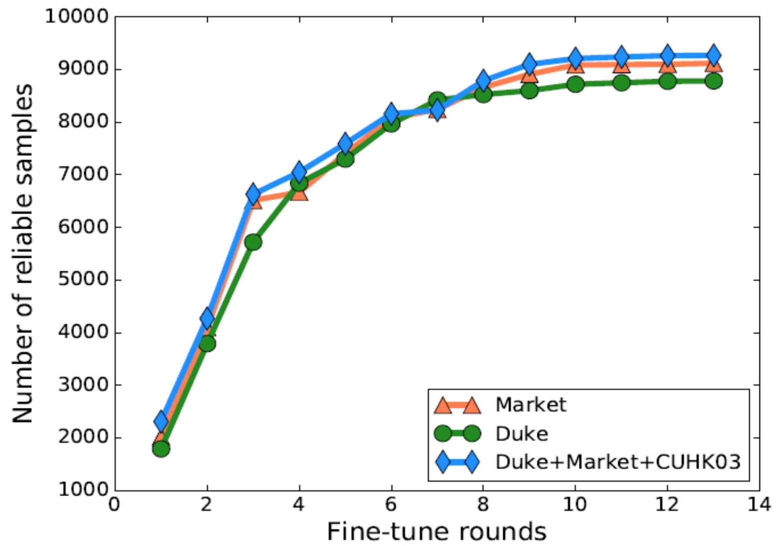
batch size =16; learning rate = 0.001; training 13 epochs

Evaluation

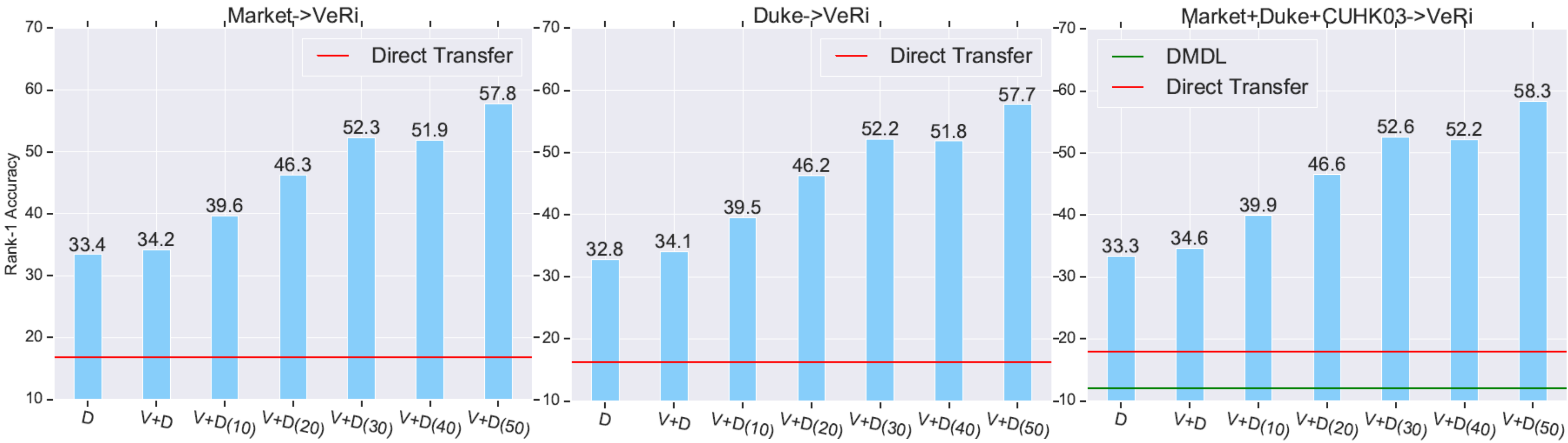


A small crop of vehicle clustering results for visualization

More reliable samples are be selected to fine-tune the CNN model



Evaluation



e.g. V means VeGAN and D means DUPL

- Training with more samples in source domain did not improve performance
- Less diversified training samples in multiple datasets
- Learning domain invariant feature is very difficult

Evaluation

S->T	Market -> VeRi					Duke -> VeRi					Market+Duke+CUHK03 -> VeRi				
methods	rank-1	rank-5	rank-10	rank-20	mAP	rank-1	rank-5	rank-10	rank-20	mAP	rank-1	rank-5	rank-10	rank-20	mAP
UMDL [17]	–	–	–	–	–	–	–	–	–	–	12.1	25.1	33.4	56.9	3.2
DT	16.8	28.7	36.8	58.8	5.3	16.3	27.9	35.8	58.1	5.2	17.9	29.2	37.4	59.1	5.4
DUPL	33.4	48.2	56.2	65.3	9.9	32.8	48.1	56.1	65.2	9.7	33.3	48.1	56.1	65.2	9.8
VeGAN+DUPL	34.2	49.1	57.3	66.4	11.3	34.1	48.9	57.2	66.2	11.2	34.6	49.1	57.2	66.5	11.4
VeGAN+DUPL(10)	39.6	55.1	62.8	71.1	14.2	39.5	54.9	62.7	70.8	14.1	39.9	55.2	62.7	71.2	14.3
VeGAN+DUPL(20)	46.3	61.2	68.5	75.8	16.9	46.2	60.9	68.4	75.5	16.8	46.6	61.3	68.4	75.9	17.1
VeGAN+DUPL(30)	52.3	66.3	73.2	80.5	19.1	52.2	66.1	73.1	80.2	19.0	52.6	66.4	73.1	80.6	19.3
VeGAN+DUPL(40)	51.9	67.8	74.7	81.6	19.8	51.8	67.6	74.6	81.3	19.8	52.2	67.9	74.6	81.7	20.0
VeGAN+DUPL(50)	**57.8**	**71.9**	**78.7**	**84.6**	**21.9**	**57.7**	**71.7**	**78.6**	**84.3**	**21.8**	**58.3**	**72.1**	**78.7**	**84.6**	**22.1**

Improvement of +18.1% in rank-1 accuracy trained on Market

- Using more IDs in target domain as training sample is always beneficial
- At a higher cost of labeled expenses
- SOTA performance in unsupervised and semi-supervised setting

Conclusion

Contribution:

- Propose a novel transfer learning model **VeGAN** for image-image translation
- Implement a simple but effective form of distant domain adaptation method **DUPL**

Future work:

- Explore the DUPL method to handle multiple target domains
- Make model more robust to unknown target domain with litter training datasets

Thank you!