Pseudo Graph Convolutional Network for Vehicle ReID

Wen Qian
qianwen2018@ia.ac.cn
Institute of Automation, Chinese Academy of Sciences
University of Chinese Academy of Sciences
Beijing, China

Zhiqun He
hezhiqun@sensetime.com
SenseTime Group Limited
Beijing, China

Silong Peng
silong.peng@ia.ac.cn
Institute of Automation, Chinese Academy of Sciences
Beijing, China

Chen Chen*
chen.chen@ia.ac.cn
Institute of Automation, Chinese Academy of Sciences
Beijing, China

Wei Wu
wuwei@sensetime.com
SenseTime Group Limited
Beijing, China

ABSTRACT

Image-based Vehicle ReID methods have suffered from limited information caused by viewpoints, illumination, and occlusion as they usually use a single image as input. Graph convolutional methods (GCN) can alleviate the aforementioned problem by aggregating neighbor samples' information to enhance the feature representation. However, it's uneconomical and computational for the inference processes of GCN-based methods since they need to iterate over all samples for searching the neighbor nodes. In this paper, we propose the first Pseudo-GCN Vehicle ReID method (PGVR) which enables a CNN-based module to performs competitively to GCN-based methods and has a faster and lightweight inference process. To enable the Pseudo-GCN mechanism, a two-branch network and a graph-based knowledge distillation are proposed. The two-branch network consists of a CNN-based student branch and a GCN-based teacher branch. The GCN-based teacher branch adopts a ReID-based GCN to learn the topological optimization ability under the supervision of ReID tasks during training time. Moreover, the graph-based knowledge distillation explicitly transfers the topological optimization ability from the teacher branch to the student branch which acknowledges all nodes. We evaluate our proposed method PGVR on three mainstream Vehicle ReID benchmarks and demonstrate that PGVR achieves state-of-the-art performance.

CCS CONCEPTS

• Computing methodologies → Computer vision; **Neural networks**; **• Information systems → Retrieval models and ranking**.

*Corresponding author.

MM '21, October 20–24, 2021, Virtual Event, China

ACM ISBN 978-1-4503-8651-7/21/10…$15.00
https://doi.org/10.1145/3474085.3475462

KEYWORDS

Vehicle ReID, Graph Convolutional Network, Knowledge Distillation, Pseudo-GCN Vehicle ReID, Distillation Evaluation Metric

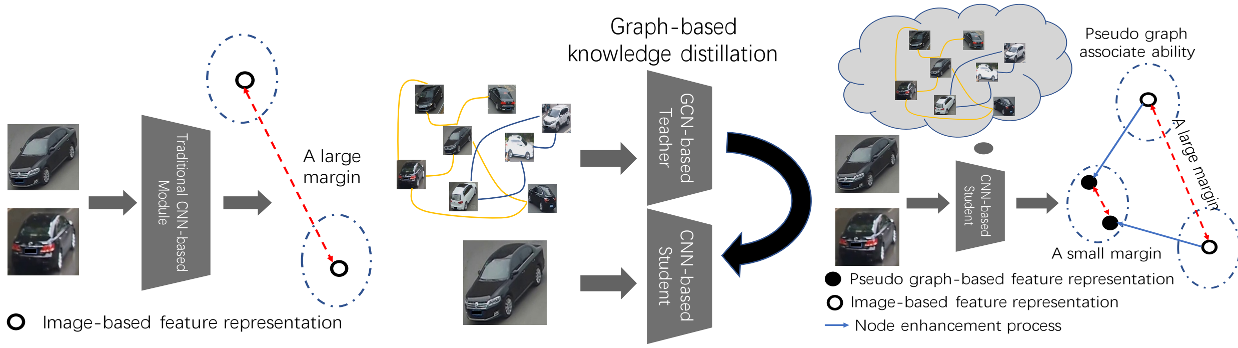
ACM Reference Format:
Wen Qian, Zhiqun He, Silong Peng, Chen Chen, and Wei Wu. 2021. Pseudo Graph Convolutional Network for Vehicle ReID. In *Proceedings of the 29th ACM International Conference on Multimedia (MM '21), October 20–24, 2021, Virtual Event, China.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3474085.3475462

1 INTRODUCTION

Most existing Vehicle ReID methods employ a single image as input which provides limited information due to viewpoints, illumination, and occlusion [2, 5, 11], and results in performance loss. As shown in Fig. 1(a), two single images of the same vehicle with different views have a large difference in appearance which leads to a large margin between their final feature representations (image-based feature representation). Graph convolutional network [9, 21] can strengthen the representation ability of the image-based feature by aggregating the information from neighbor nodes. Such graph-based information is more suitable for alleviating the aforementioned limitations in Vehicle ReID than image-based representation.

Some researchers employ a graph convolutional networks (GCN) to cooperate with ReID methods [3, 16, 41], and show that the combination of GCN and ReID tasks can achieve better performance. Despite the encouraging progress, existing GCN-based ReID methods need to cooperate with neighbors' information during both the training time and inference time, and here comes two limitations for deploying GCN-based methods in real-life scenarios: for the first one, we can't obtain valid neighbor information in some real-life scenarios such as the urban traffic escape scenario which usually only has a snapshot; for the other one, large-scale ReID scenarios usually have a large gallery set and the iteration process over them for searching neighbor nodes costs much time.

In this paper, the Pseudo-GCN Vehicle ReID method (PGVR) is proposed to capacitate a pure CNN-based module to perform like a GCN-based module during the inference time. Our proposed PGVR consists of a two-branch network that contains a GCN-based teacher branch and a CNN-based student branch, and a graph-based



(a) The process of image-based method (b) The training process of Pseudo-GCN method (c) The inference process of Pseudo-GCN method

Figure 1: A comparison between the image-based method and the Pseudo-GCN-based method which shows Pseudo-GCN performs better with a lightweight inference process. The training and inference processes of the image-based method are the same while the Pseudo-GCN-based method has different training and inference processes.

knowledge distillation for information transferring. The teacher branch aims to learn a robust feature representation with a stripe-based specification and a ReID-based GCN module. The stripe-based specification is inspired by PCB [38], which proves that a coarse local specification can lead to considerable boosting in performance. Moreover, the student branch mainly focuses on the global features with a pure CNN-based module. The kernel of our method is graph-based distillation which transfers the topological optimization ability from the teacher branch (GCN-based module) to the student branch (CNN-based module). By graph-based knowledge distillation shown in Fig. 1, the CNN-based student branch has a pseudo-graph associate ability which drives the CNN-based module to perform competitively to the teacher module and thus is named as the Pseudo-GCN-based module.

For accommodating the GCN module in the teacher branch to Vehicle ReID, the GCN module works under the supervision of the Vehicle ReID losses. The Vehicle ReID losses optimize our ReID-based GCN (RGCN) in both classifier-space and metric-space, and a more precise topological relation can be provided for the student branch to learn from. Specifically, we construct an anchor-based subgraph (AS) that holds two-level neighbors for a more comprehensive relationship in RGCN. Different from traditional knowledge distillation, our proposed graph-based knowledge distillation (Graph-KD) aims to transfer graph and logic information which is consistent with classifier-space and metric-space information learned by RGCN. Given the anchor-based subgraphs constructed from the teacher and the student branch, we argue that they only emphasize neighbor nodes while neglect that the student and teacher branch tends to have different neighbor nodes. The inconsistency between the node set may lead to negative knowledge transfer. To address the aforementioned limitation, we enlarge the AS to a broaden anchor-based subgraph (BAS) by taking all nodes into the node-set.

We evaluate the performance of Pseudo-GCN Vehicle ReID (PGVR) on three popular Vehicle ReID benchmarks i.e., VeRi776 [26], VehicleID [27], and VERI-WILD [28], and propose a new evaluation metric to evaluate the performance gain from distillation methods. Experiments show that our method achieves the best performance among all the compared Vehicle ReID methods, and Pseudo-GCN-based methods can be widely used in many real senses by driving CNN-based methods to perform like GCN-based methods.

The contributions of this paper are:

• We introduce the first Pseudo-GCN Vehicle ReID method (PGVR) by which a CNN-based module obtains competitive performance to GCN-based module. PGVR alleviates the cost of time and computation for searching the neighbor nodes in GCN-based methods during inference.

• A graph-based knowledge distillation (Graph-KD) is proposed to provide a ReID style knowledge transfer from both classifier-space and metric-space which takes all nodes for aligning the gaps between student and teacher branches.

• Our PGVR has been evaluated on three popular benchmarks and outperforms all other methods consistently. Specifically, we further propose a distillation evaluation metric, and Graph-KD performs better than other distillation methods based on this metric.

2 RELATED WORK

2.1 Vehicle ReID

Although Vehicle ReID has achieved great progress [4, 13, 38], there still exists challenges such as the inter-class similarity and the inner-class difference. Several Vehicle ReID methods are proposed to solve these problems, and can be classified as local-based methods, attribute-based methods, and viewpoint-aware methods.

Firstly, the local-based methods attempt to specify some important regions which can cooperate well with existing Vehicle ReID methods. Unsupervised specification methods usually adopt a multi-branch structure to extract coarse local information by dividing the shared feature into multiple parts [2, 4, 58]. The supervised specification methods [11, 53] detect the locations of specific parts (lights, license plates, logos). Secondly, the attribute-based methods use the semantic and identifiable information contained in attributes [17, 22, 25, 36, 37, 48] to boost the performance of Vehicle ReID. In [24], researchers intend to train the attribute classification task and the Vehicle Re-ID task simultaneously by a joint learning approach. Finally, viewpoint changes can cause a large variety of intra-class differences in vehicle Re-ID, and many methods are proposed and focus on solving this problem [5, 29].

2.2 Graph Neural Networks

The concept of graph neural network (GNN) [34] was firstly proposed before 2010, and has been proved to be an effective model for non-grid data which is usually represented by a set of nodes along with a graph that denotes the relationship among nodes [21, 45, 54]. As one of the most influential GNN models, Graph Convolutional Network (GCN) [6, 7, 9, 21] targets graph-structured data through layer-wise propagation of node features. Many GCN-based methods are proposed: GAT [40] employs the attention mechanism to cooperate with GCN, and makes it possible to learn the weight for each neighbor automatically; GraphSAGE [10] samples the neighbors rather than using all of them which makes the GCN model scalable for huge graph; researchers adopt an instance pivot subgraph algorithm to find its 1-hop and 2-hop neighbors which leads to a more reasonable GCN model [43]. Many other works incorporate traditional prediction mechanisms, i.e., label propagation to improve the performance of GNN. UniMP [35] uses a shared message-passing network to fuse feature aggregation and label propagation.

2.3 Knowledge Distillation

Knowledge distillation is first proposed by Hinton in [15] which transfers knowledge from a teacher model to a student model and thus the student model can obtain a similar performance as the teacher model. Besides the soft-predictions from the teacher model, the intermediate activation [49, 57], the relation between samples [30, 31, 44] in a graph can also be utilized for knowledge transfer. Researchers transfer the attention instead of the feature itself to get a better distillation performance in [49]; Wu et al. propose to transfer a similarity matrix and provide a Re-ID model integration system that can be incrementally learned [44]. Here are also a few works combining knowledge distillation with GNN, and their network structure and motivation are quite different from ours. Yang et al. propose to use distillation in GCN for compressing a large GCN model to a shallow one with fewer parameters [47]; Reliable Data Distillation (RDD) [8, 50] trains multiple GCN students with the same architecture and then concentrates them for better performance; Graph Markov Neural Networks (GMNN) [33] employs two GCNs with different reception sizes to learn from each other which can be viewed as a knowledge distillation method. Compared with the aforementioned methods, the graph-based knowledge distillation in our method aims to distill from graph-based branch to CNN-based branch and enable the CNN-based student branch to perform competitively to the GCN-based branch.

3 METHODOLOGY

In this section, we will start by introducing the notations and formalizing the adopted ReID-based graph convolutional network used in the Pseudo-GCN Vehicle ReID method (PGVR). Then we will propose the architecture of PGVR which consists of a GCN-based teacher branch and a CNN-based student branch. Afterward, we propose graph-based knowledge distillation (Graph-KD) which provides a ReID style knowledge transfer from both classification-space and metric-space and aligns the node-sets between student and teacher branches.

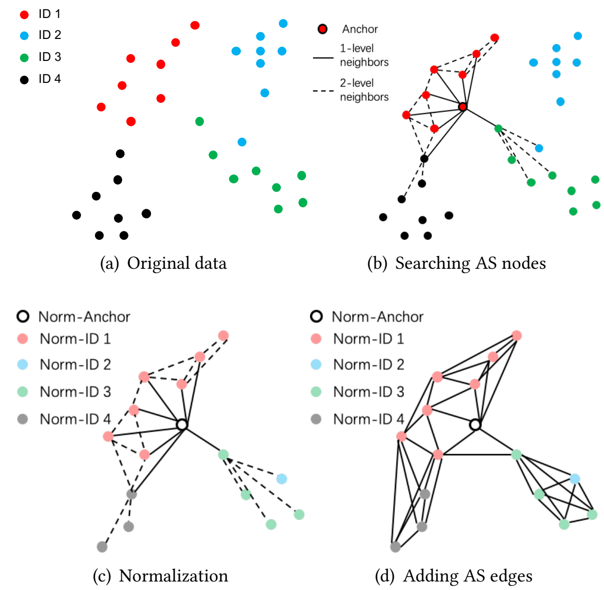


Figure 2: The construction of anchor-based subgraph for a given sample batch: a) the original batch of samples have 4 vehicles and 8 samples for each vehicle; b) search AS nodes and set the 1-level node number as 8 and 2-level node number as 4; c) normalizing the node feature and removing the no-neighbor nodes; d) assigning edges to the 4-NN nodes in AS nodes.

3.1 ReID-based Graph Convolutional Network

3.1.1 Notations. Assuming that we have a collection of vehicle features $X = [x_1, x_2, ..., x_N]^T \in R^{N \times D}$ in a batch, where N is the number of vehicles, D is the dimension of the features, and each feature x is treated as a node v in our ReID-based graph convolutional network (RGCN). For generality, we use $G(V, E)$ to describe the graph of our network, where V and E represent the node-set and edge-set respectively. The corresponding node feature can be represent as $F = x_q | q \in V$ and $F \in R^{M \times D}$, where $M = len(V)$ represents the number of nodes in V.

3.1.2 Anchor-based Subgraph. ReID tasks usually employ a triplet loss to optimize the features from metric space and organize the samples at a $b \times c$ mode which means that b vehicles each with c samples in a batch. The above sample mechanism enables for each anchor sample, there exist $c - 1$ samples with the same Vehicle-ID in a batch. We build an anchor-based subgraph (AS) to describe the relationship among nodes, which is used to estimate the linkage possibility between the anchor node and its neighbor nodes. Given the vehicle features X, we generate the AS follow the next steps: firstly, the two-level nearest neighbor node-sets of each anchor node are located; then we normalize the node features and add edges among nodes. An illustration of anchor-based subgraph (AS) can be seen in Fig. 2.

Step 1: Nodes searching. For each node p in vehicle features $X = [x_1, x_2, ..., x_N]^T \in R^{N \times D}$, we search for its two-level nearest neighbors. The number of the picked neighbor number $k_i, i = 1, 2$ in each level are vary and set followed the description in Sec.4.2.2. For

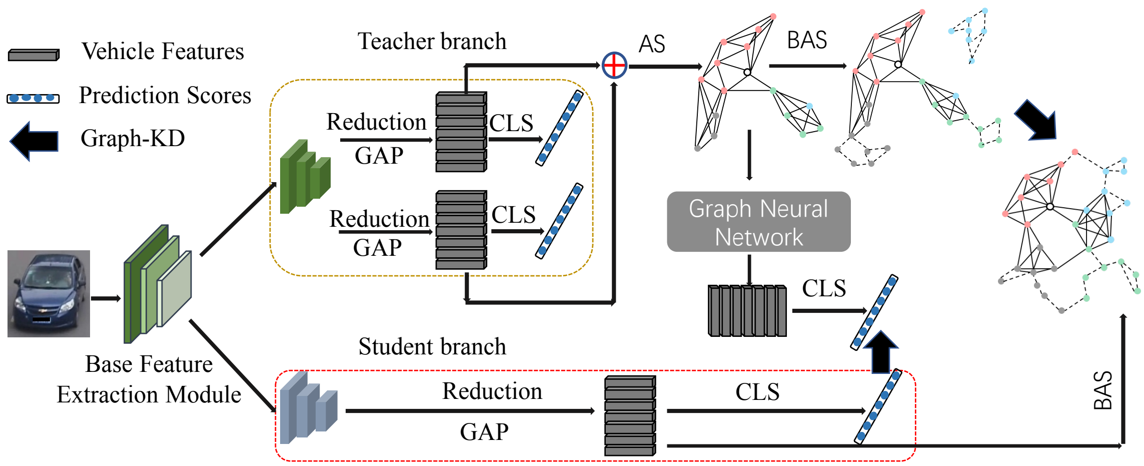


Figure 3: The network architecture of our proposed Pseudo-GCN Vehicle ReID (PGVR). The student branch and teacher branch have a shared backbone for base feature extraction. And both branches need to calculate the broaden anchor-based subgraph for knowledge transfer in Graph-KD.

example in VeRi-776 (16 vehicles, 8 samples for each vehicle), the node set V_p with $k_1 = 8$ and $k_2 = 4$ consists 8 nearest neighbors of p, 4 nearest neighbors of each 1st level node. The intention is to let the high-order neighbors in V_p can provide auxiliary information of the context between an anchor and its closest neighbor. Even though the first level neighbors can describe the most important neighbor nodes of an anchor ideally, we still employ the second-level nodes as a supplementary since sometimes a node q with the same vehicle of p will be classified outside the first-level nodes.

Step 2: Normalization and Edge-Connection. After obtaining the AS node-set V_p of an anchor node p and their node features $X_p : \{x_i | i \in V_p\}$, we conduct a subtract operation for encoding the anchor node feature into the node features of AS:

$$F_p = [..., x_p - x_q, ...]^T, for\ all\ q \in V_p, \tag{1}$$

where $F_p \in R^{|V_p \times D|}$ represents the normalized node features.

Then, we add edge connection among the node-set V_p. For each node q in V_p, we find the top m nearest neighbors among all 128 samples in a batch firstly. If a node r of mNNs appears in V_p, an edge connection $A_p(q, r) = 1$ is added to the edge set E_p. Finally, the topological structure of AS can be represented by an adjacency matrix $A_p \in R^{|V_p| \times |V_p|}$ and the node feature matrix F_p.

3.1.3 Graph Convolution Network for Vehicle ReID. We propose a ReID-based graph convolution network (RGCN) modified from GCN[2] under the supervision of the ReID task (triplet loss and Vehicle-ID loss). Moreover, the aforementioned anchor-based subgraph contains the context A_p and F_p of the anchor node which is highly valuable for determining whether the nodes belong to the same vehicle-ID. So the GCN layers in RGCN take the node feature matrix $F_p \in R^{N \times d_{in}}$ and the adjacency matrix A_p as input, and output an aggregated feature matrix $Y_p \in R^{N \times d_{out}}$, where N denotes the number of nodes in V_p, d_{in}, d_{out} are the dimension of F_p and Y_p. Formally, the graph convolution layers can be formulated as following:

$$\begin{aligned} Y_p &= \sigma([F_p || GF_p]W) \\ G &= g(F_p, A_p), \end{aligned} \tag{2}$$

where $\sigma(\cdot)$ is the non-linear activation function. The aggregation matrix $G \in R^{N \times N}$ sums each row up to 1, and $g(\cdot)$ is a function of F_p and A_p such as mean aggregation, weighted aggregation, or attention aggregation. Operator || represents matrix concatenation along the feature dimension, and W of size is $2d_{in} \times d_{out}$ is a weight matrix of the GNN layer which can be learned.

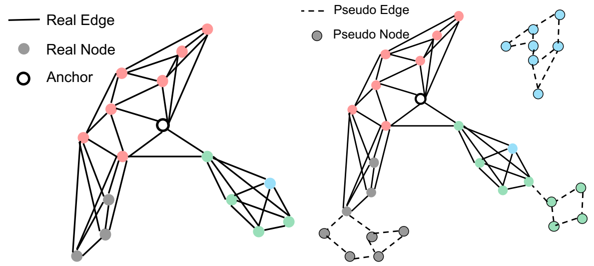
Our ReID-based Graph convolution Network employs four RGCN layers with the output dimension of these RGCN layers drops from 2048 to 512 and then rises from 512 to 2048 to keep the output feature the same dimension as the ReID feature. Previous GCN-based methods are only supervised by a two-class classification task that neglects the node-relation in the anchor-based subgraph. Moreover, we use ReID losses, i.e., triplet loss and Vehicle-ID loss as the supervisions for the training process of RGCN which provide more difficult samples for both classification-space and metric-space. Specifically, the employment of triplet loss on V_p which only takes the neighbor nodes (the most similar nodes) into consideration pushes the RGCN model to perform better on Vehicle ReID tasks, since the most similar nodes V_p is difficult to distinguish in feature metric space.

3.2 Pseudo-GCN Vehicle ReID Method

In this subsection, we first introduce the network architecture of the Pseudo-GCN Vehicle ReID method (PGVR). Then we introduce the graph-based knowledge distillation which is specified for graph-based ReID scenes. Finally, we introduce the training details of our network.

3.2.1 Network . For an effective Pseudo-GCN Vehicle ReID method, we employ a two-branch network architecture including a GCN-based teacher branch and a CNN-based student branch. In this part, we will introduce the motivations and the design for these two branches, which can be seen in Fig. 1 and Fig. 3.

The teacher branch employs a local-specification mechanism for mining more local information which is a classic mode first proposed in [38]. The combination of local-specification mechanism and GCN can lead to a performance-boosting in the teacher branch, and more knowledge can be transferred to the student branch. For



(a) The anchor-based subgraph (b) Broaden anchor-based subgraph

Figure 4: The process of broaden anchor-based subgraph.

a given sample G, we firstly extract a basic feature representation g by ResNet50 [12], and then the basic feature representation g will be divided horizontally into two parts: the upper one g^u and the lower one g^b. This kind of division matches the architecture of vehicles well since the upper one usually contains information such as windows or drivers while the lower one usually focuses more on the plates or the wheels. The divided features g^u and g^b will be pooled by global average pooling and then get local specification features x^u and x^b. On the one hand, both the local specification features x^u and x^b are supervised by triplet loss and vehicle-ID loss at the common ReID mode respectively; on the other hand, the local specification features x^u and x^b are concentrated as $X = [x_1, x_2, ..., x_N]^T \in R^{N \times D}$ and then be used as the input of ReID-based graph convolution network (RGCN). It needs to be claimed that the ReID losses adopted for RGCN are operating under a stricter sample pair and can reach a more precise performance.

The student branch is organized at the mode of common ReID methods which employs a ResNet to extract the base feature. Then the global feature will be obtained under the operation of global average pooling which is supervised by the losses of ReID tasks.

3.2.2 Graph-based Knowledge Distillation. As aforementioned, we propose a ReID-based graph neural network strengthen the feature representation in the teacher branch through aggregating neighbor nodes' information, and drive the CNN-based student model to have a competitive performance through efficient graph-based knowledge transfer (Graph-KD). Unlike some previous methods [46] which distills knowledge from a large GCN module to a small GCN module, Graph-KD distills knowledge from GCN-based module to CNN-based module. Moreover, Graph-KD works under the guidance of Vehicle ReID tasks which transfers two kinds of information including classification information and feature metric information.

Assuming that the output vehicle feature from student branch and teacher branch as $X^s = [x_1^s, x_2^s, ..., x_N^s]^T$ and $X^t = [x_1^t, x_2^t, ..., x_N^t]^T$. Then we need to generate the corresponding subgraphs $G_t(V_t, E_t)$ and $G_s(V_s, E_s)$ for the student branch and the teacher branch based on the input of X^s and X^t. Finally, we can get a corresponding feature set $F = [f_1, f_2, ..., f_M]^T$ for each anchor node, where M denotes the number of nodes contained in V. Two prediction score S_s and S_t are obtained from the corresponding classifier layer during the training period, where S_s is directly obtained from X^s. Before obtaining the teacher prediction score S_t, we need to pick out the center feature in each F and concentrate them as the input of the classifier layer. The first item in Graph-KD is Kullback-Leibler Divergence distillation proposed by Hinton, and we use it to distill the classifier prediction score S_s and S_t:

$$KL(S_s||S_t) = \sum S_s log \frac{S_s}{S_t}. \tag{3}$$

As it's known to all, the most important capacity of GCN is the feature aggregation ability come from topological optimization (the teacher branch), and we aim to enable the CNN-based model (the student branch) to have a competitive performance in topological optimization through graph-based knowledge distillation. So another item of Graph-KD is GCN-CNN distillation which aims to transfer the graph information from the teacher branch to the student branch. We obtain the corresponding subgraphs $G_t(V_t, E_t)$ and $G_s(V_s, E_s)$ in both branches to describe the nodes relationship. However, we find that the node-set V_t and V_s tend to be different since different states and architectures of two branches, and directly distill on $G_t(V_t, E_t)$ and $G_s(V_s, E_s)$ may do harm to the ReID performance. To align the node-set in two branches, we propose a broaden anchor-based subgraph as shown in Fig. 4. The broaden anchor-based subgraph (BAS) takes all nodes into consideration. The non-neighbor nodes outside V in the sample batch are termed as pseudo-node and we add pseudo-edges among the pseudo-nodes and the uNN closest nodes. The formula for GCN-CNN distillation is:

$$L1(G_s^B||G_t^B) = \sum_{i=1,j=1}^{N} |G_s^B(i,j) - G_t^B(i,j)| \tag{4}$$

where G_s^B and G_t^B denote the broaden anchor-based subgraph, i and j represent for each element in the corresponding broaden anchor-based subgraph, and N is the batch-size. The edge connections we used in $G(V, E)$ are binary, so L_1 or L_2 loss works at the same procedure.

We have to point out that the real feature aggregation or topological optimization should all rely on neighbor nodes, so an image-based CNN module can only be named as Pseudo-GCN which works in a quite different mode. The information transferred in Pseudo-GCN indeed is a fine-grained ReID ability and a feature alignment ability that can be learned from a GCN-based teacher module. Finally, the equation for structural graph-based knowledge distillation can be described as:

$$L_{Graph-KD} = KL(S_s||S_t) + L1(G_s^B||G_t^B) \tag{5}$$

3.2.3 Training Details. Our proposed PGVR can be separated into two parts during training time: a GCN-part and a CNN-part, and both the two parts are trained at an end-to-end mode. For a given sample batch, we first calculate corresponding feature representations X^s and X^t in both student branch and teacher branch, and a classifier layer is used to calculate prediction scores S_s and S_t. And then we obtain anchor-based subgraph (AS) G_s and G_t, broaden anchor-based subgraph (BAS) G_s^B and G_t^B, RGCN's output features X_T^O, RGCN's output prediction S_T^O from vehicle features X^t and X^t. Firstly, we calculate GCN loss with RGCN's output X_T^O and S_T^O:

$$L_{gcn} = Triplet(X_T^O, L) + ID(Score_T^O, L), \tag{6}$$

where L represents for the original labels of the input sample batch.

Table 1: **Comparison with state-of-the-arts. It includes mAP and CMC@1 on VeRi-776; CMC@1 on three test sets of small, medium, and large on VehicleID; mAP on three test sets of small, medium, and large on VERI-WILD. For the three test sets on VehicleID and VERI-WILD, they are represented by S, M, and L respectively. Finally, "/" indicates missing parts of the experiments, ".n.d" represents the methods without distillation.**

	Method	VeRi-776		VehicleID			VERI-WILD		
	(%)	mAP	CMC@1	CMC@1 (S)	CMC@1 (M)	CMC@1 (L)	mAP (S)	mAP (M)	mAP (L)
Part-based	PGAN [51]	79.3	96.5	77.8	/	/	74.1	/	/
	PRN [11]	74.3	94.3	78.4	75.0	74.2	/	/	/
	PVEN [29]	79.5	95.6	84.7	80.6	77.8	82.5	77.0	69.7
	GLAMOR [39]	80.3	96.5	78.6	/	/	77.2	/	/
Attribute-based	AGNet-ASL [42]	71.59	95.61	71.15	69.23	65.74	/	/	/
	DJDL [24]	/	/	78.6	74.7	72.0	/	/	/
	XG-6-sub-multi [53]	/	/	76.1	73.1	71.2	/	/	/
	SAN [32]	72.5	93.3	79.7	78.4	75.6	/	/	/
Attention-based	AAVER [18]	61.2	89.0	74.7	68.6	63.5	/	/	/
	SEVER [19]	79.6	96.4	79.9	77.6	75.3	83.4	78.7	71.3
Others	GSTE [1]	59.4	/	87.1	82.1	79.8	/	/	/
	VAMI [56]	61.3	89.5	63.1	52.9	47.3	/	/	/
	DCDLearn [59]	70.4	92.8	82.9	78.7	75.9	/	/	/
Ours	Baseline (PGVR.n.d)	80.8	96.7	82.9	80.1	78.3	83.5	79.0	74.6
	PGVR	83.8	97.3	87.0	82.6	80.5	84.6	80.2	75.0

After we optimize the GCN-part of PGVR by L_{gcn}, we can calculate the CNN-based losses:

$$L_{cnn} = \sum_{i=s \ or \ t} Triplet(X_i, L) + ID(S_i, L) + L_{Graph-KD}, \quad (7)$$

and L_{cnn} is used to optimize the CNN-part of PGVR.

4 EXPERIMENTS

We detail the implementation and evaluation of PGVR in this section. The datasets and evaluation metrics are introduced in section 4.1; the implementation details are introduced in section 4.2; the comparisons with State-of-the-Arts and other distillation methods are introduced in section 4.3; the ablation study and evaluation based on PGVR are introduced in section 4.4 finally.

4.1 Datasets and Evaluation Metrics

Dataset We evaluate our Pseudo-GCN Vehicle ReID method on three popular Vehicle ReID benchmarks: VeRi776 [26], VehicleID [27], and VERI-WILD [28]. VeRi776 [26] is a classic Vehicle ReID benchmark, which contains 776 identities collected by 20 cameras in a real-world environment. The whole dataset is split into 576 vehicles for training and 200 vehicles for testing. VehicleID [27] is a large-scale dataset collected by multiple cameras during the daytime on the open road, which contains 26,267 vehicles and 221,763 images in total. VERI-WILD [28] is another large-scale dataset, and it consists of 40,671 vehicles and 416,314 images. Moreover, VERI-WILD [28] is collected by 174 cameras during a month which is a long period.

Evaluation Metrics For the ReID task, we follow the same evaluation protocols which are used in previous works [27, 52, 55]. Mean Average Precision (mAP) and the cumulative matching characteristics at Rank1 (CMC@1) are employed to evaluate the performance of PGVR. The value of CMC@1 represents the chance of the correct match appearing in the top 1 of the ranked candidate list. Moreover, it should be noticed that the VehicleID [27] benchmark pays more attention to CMC@1 for its unique test sets.

For the distillation task, we propose Student-Teacher Transfer Efficiency (STTE) to measure the efficiency of knowledge distillation among teacher model and student model:

$$STTE = \frac{S_{mAP}.n.d - T_{mAP}}{S_{mAP} - T_{mAP}}, \quad (8)$$

where $S_{mAP}.n.d$ represents the mAP of student model without knowledge distillation, S_{mAP} and T_{mAP} represent the mAP of models with distillation. The $STTE$ quantifies the improvements in the student model brought from knowledge distillation, and mAP can be replaced by other accuracy metrics. A perfect knowledge distillation can reduce the gap between the student model and teacher model corresponds to $STTE$ equal to 1.

4.2 Implementation Details

4.2.1 Experimental Setting. We perform our experiments in PyTorch on a machine with 8 NVIDIA Titan Xp GPU. The visible images are resized to $256 \times 256 \times 3$ for both model training and testing, and a common erasing and a flipping operation are employed for data augmentation. The CNN-part and GCN-part of PGVR are trained together at an end-to-end mode. We employ two independent optimizers which use the same parameters to optimize the CNN-part and GCN-part respectively. The optimizers used in our method is Adam with the weight decay factor of 0.0001, and the learning rate is initialized to 1e-4 and decreased by a factor of 0.1 after the 40^{th} and 70^{th} epoch.

4.2.2 Parameter Selection. Several hyper-parameters need to be chosen for the construction process of AS and BAS: the number of picked nearest nodes $k_i,\ i = 1, 2$ in the first and second level, the number of mNN nodes for edge connection in AS, and the

Table 2: Comparison with other knowledge distillation methods on VeRi-776.

Approaches (%)	Distillation method	SMAP.n.d	SMAP	TMAP	STTE
KD [15]	Soft Lables	80.7	82.8	83.2	84.0
GKD [23]	Soft Lables	80.7	82.2	83.2	60.0
CKD [57]	Attention	80.7	81.3	83.2	24.0
AKD [49]	Attention	80.7	81.9	83.2	48.0
FTKD [20]	Attention	80.7	82.0	83.2	52.0
ABKD [14]	Attention	80.7	82.2	83.2	60.0
RKD [30]	Relation	80.7	81.5	83.2	32.0
CCKD [31]	Relation	80.7	82.4	83.2	68.0
Our Graph-KD	Graph	80.7	83.8	84.0	94.0

Table 3: Ablation study of different picked number k_1, k_2, and m in PGVR (VeRi-776)

Scheme	Method	k_1	k_2	m	MAP (%)	CMC@1 (%)
a	baseline	/	/	/	80.7	95.6
b	PGVR	4	4	4	83.3	96.9
c	PGVR	8	4	4	83.8	97.3
d	PGVR	12	4	4	82.7	96.6
e	PGVR	8	2	4	83.2	96.9
f	PGVR	8	6	4	82.3	96.3
g	PGVR	8	8	4	81.5	95.8
h	PGVR	8	4	2	81.9	96.1
i	PGVR	8	4	6	83.4	97.0

Table 4: Ablation study on the number of layers in RGCN(VeRi-776)

Scheme	Method	Layer Number	MAP (%)	CMC@1 (%)
a	baseline	/	80.7	95.6
b	PGVR	2	83.3	96.9
c	PGVR	4	83.8	97.3
d	PGVR	8	84.2	97.2

number of uNN nodes for edge construction of pseudo-nodes. As we claimed in the previous section: k_1 should be set consistently with the instance number for each vehicle during training, i.e., 8 for VeRi-776 and 4 for VehicleID and VERI-WILD. Different values of k_2 have experimented while we find that $k_2 > k_1$ does not bring any performance gain, so $k_2 = 4$ is set for all datasets. Then we explore the impact of different values of m and u and find that $m > u$ can achieve a good performance and we set m and u as 4 and 2.

4.3 Comparisons with Other Works

4.3.1 Comparisons with State-of-the-Arts. We compare PGVR with a wide range of state-of-the-arts Vehicle ReID methods, including (1) part-based approaches: PGAN [51], PRN [11], PVEN [29], and GLAMOR [39]; (2) attribute-based approaches: AGNet-ASL [42], DJDL [24], XG-6-sub-multi [53], and SAN [32]; (3) attention-based approaches: AAVER [18] and SEVER [19]; (4) other interesting approaches: GSTE [1], VAMI [56], and DCDLearn [59].

The comparisons are shown in Table 1, and all our experiment results shown in Table 1 are all tested on the student branch:

1) Our baseline already achieves state-of-the-art performance on most benchmarks by adopting a shared backbone which can also be improved by the GCN-based teacher. PGVR still achieves new gains and outperforms other competitors by a larger margin.

2) Compared to the part-based and attribute-based methods [11, 29, 32, 53], PGVR achieves significant improvement, e.g. up to 3.5% mAP on VeRi-776, 3.3% mAP on VehicleID, and 2.1% mAP on VERI-WILD, which validates that the student in PGVR can learn better fine-grained information than these methods.

3) Compared to the attention-based methods [17, 18], PGVR achieves significant improvement, e.g. up to 4.2%mAP on VeRi-776, 7.1% mAP on VehicleID, and 1.2% mAP on VERI-WILD.

4.3.2 Comparison with other distillation methods. Although we claim that our Graph-KD enables the CNN-based student branch to perform competitively to the teacher branch. There also exists many previous knowledge distillation methods such as knowledge distillation (KD) [15], guided knowledge distillation (GKD) [23], relational knowledge distillation (RKD) [30], Channel-Wise attention for knowledge distillation (CKD) [57], correlation congruence for knowledge distillation (CCKD) [31], factor transfer (FTKD) [20], Paying More Attention to Attention (AKD) [49] and activation boundaries knowledge distillation (ABKD) [14]. We compare our graph-based knowledge distillation with these distillation methods on the ReID task (VeRi-776) as shown in Table 2. Experiments show that our method achieves 94% in STTE and 83.8% in mAP which both outperform all other distillation methods.

4.4 Ablation Study and Evaluation

4.4.1 The impact of picked number k_1, k_2, and m in AS. Table 3 shows the influence of the picked nodes number k_1, k_2, and m in AS on VeRi-776. We conduct 8 schemes of experiments on these three parameters, and get the following conclusions:

1) From Table 3 (b)(c)(d), we find that when k_1 is set as the number of instance in the benchmark, i.e., set as 8 on VeRi-776 , it works better. If k_1 is smaller than 8, the GCN module can't iterate all nodes with the same Vehicle-ID which will lead to a slight performance decrease (As shown in Table 3 (b)), but when k_1 is larger than 8, the anchor node will aggregate some false information and nodes belong to different Vehicle-ID will become hard to distinguish (As shown in Table 3 (d)).

2) From Table 3 (c)(e)(f)(g), we experiment on how many nodes should be picked for k_2 in the second level and get the conclusion that $k_2 = 4$ performs better finally. When we set $k_2 = 8$, too many connections are added among nodes in AS which introduces many interference information to the original feature; and $k_2 = 2$ results in a smaller anchor-based subgraph which leads to a weaker triplet loss since too few nodes are contained.

3) From Table 3 (c)(h)(i), we also find $m = 4$ performs better and drives the nodes in the final graph distinguishable. When we increase the final connection node from $m = 4$ to $m = 6$, only a few extra nodes are added and the mAP does not drop too much. Another reason is that we only take $k_2 = 4$ second-level nodes into consideration and those nodes outside the anchor-based subgraph will not be considered even m is set to 6.

4.4.2 The impacts of graph layer number. Table 4 shows the influence of layer number in the RGCN module on VeRi-776. Benefiting

Table 5: Ablation study of the components in PGVR(VeRi-776)

Scheme (%)	C-cut	RGCN	Graph-KD	S-MAP	S-CMC@1	T-MAP	T-CMC@1
a	×	×	×	80.7	95.6	80.7	95.6
b	✓	×	×	81.3	95.8	82.0	96.3
c	✓	✓	×	82.1	96.5	83.5	96.9
d	✓	×	✓	82.3	96.5	82.2	96.4
e	✓	✓	✓	83.8	97.3	84.0	97.3

Table 6: The average time costing, number of parameters, number of flops of our method.

	Time-Costing (s)	Paras (M)	Flops (M)
Training	6.2	102.4	10210
Inferencne	0.4	54.7	8270

from the residual connection, we can extend the RGCN network deeper to obtain more powerful node representations. We compare the results from shallow to deep networks (layers=2, 4, 8), and experiments show the mAP improves from 83.3% to 84.2% in mAP with the growth of layer numbers.

4.4.3 The impact of components in PGVR. The most important components of our Pseudo-GCN Vehicle ReID method (PGVR) can be described as the coarse local-part specification mechanism (C-cut), the ReID-based graph convolutional network (RGCN), and the graph-based knowledge distillation (Graph-KD). We ablate the performance improvement brought by these parts on the VeRi-776 benchmark, and Table 5 reports the ablation results where S- and T- represent the results on the student branch and teacher branch.

1) With the employment of the coarse local-part specification mechanism (C-cut) in Table 5 (b), we find the teacher branch achieves a 1.3% gain in mAP and the student branch also achieves a 0.6% gain in mAP. We conclude that the shared backbone between two branches can also lead to some improvements by joint training.

2) In Table 5 (c), we employ the combination of the C-cut and the ReID-based graph convolutional network (RGCN), and obtain 1.4% and 2.8% gain in mAP on the student and teacher branches. This process shows that our RGCN can bring an extra improvement with no knowledge distillation, and validates its effectiveness.

3) In Table 5 (e), we examine the performance of graph-based knowledge distillation (Graph-KD), and the experiment shows that the combination of Graph-KD and RGCN can lead to a 3.1% gain in mAP on the student branch. Moreover, when we conduct Graph-KD operation without RGCN as shown in Table 5 (d), we can find that the student branch performs worse than the one with RGCN which shows Graph-KD needs to cooperate with RGCN.

4.4.4 Analysis of the Improvements in Student Branch. As can be seen in Table 1 and Table 2, we find that our Pseudo-GCN Vehicle ReID method can drive a CNN-based network to perform competitively to traditional GCN-based networks, and we attempt to explain this phenomenon in this part.

Firstly, as shown in Fig. 1, we claim that the graph-based knowledge distillation can give the student model a pseudo-graph associate ability which indeed represents the ability to understand the structure of vehicles from the process of aggregating neighbor information. So the student model can align the vehicle features better after understanding the vehicle structure which alleviates the misalignment caused by diverse viewpoints or occlusions and leads to a performance improvement.

Secondly, the performance improvement can also be explained by the fine-grained ability transferred from the teacher model. As aforementioned, the teacher model adopts a ReID-based graph convolutional network to enhance the feature representation by integrating some related information and tell the teacher branch which regions are important statistically. For example, the teacher branch may focus more on the appearance of vehicles initially, and it may emphasize more subtle features such as plates and windows after aggregating the neighbor features. Moreover, this process can be easily transferred to the student branch.

The overview of all neighbor samples with the same vehicle-ID by a RGCN-based module can enable the network to learn a feature alignment ability and a fine-grained feature extraction ability. We name the combination of the feature alignment ability and the fine-grained feature extraction ability as Pseudo-GCN in our paper.

4.4.5 Analysis of the inference speed of Pseudo-GCN-based method. As mentioned before, we claim that the Pseudo-GCN-based method only needs to deploy the global branch during inference time which is faster and economical. So we experiment on the average time costing, the number of parameters, and the number of flops for both the training process and inference process of our method in Table 6. Experiments show that Pseudo-GCN reduces the inference time from 6.2 seconds per sample to 0.4 seconds per sample, while also save 46.6% parameter costs and 20% Flops costs.

5 CONCLUSION

Given that existing GCN methods are time-costing and computational during inference, we propose a Pseudo-GCN Vehicle ReID method (PGVR) which enables a CNN-based branch to perform competitively to a GCN-based branch. We emphasize the importance of context in Vehicle ReID and propose a ReID-based graph convolutional network to aggregate the neighbor nodes' information which employ an anchor-based subgraph mechanism (AS) for graph construction. For transferring the topological optimization ability from the teacher branch to the student branch better, graph-based knowledge distillation which employs classification information and feature metric information (graph) for transferring is proposed. Moreover, we claim that the misalignment of the AS information from two branches will lead to a performance damage and adopt a broaden-anchor-based subgraph (BAS) to solve it which accounts for all nodes. We report favorably comparable results to state-of-the-art methods on popular Vehicle ReID benchmarks and show our proposed PGVR can save 93.5% costing time, 46.6% parameter costs, and 20% Flops during the inference period. Finally, we also propose an evaluation metric to evaluate the efficiency of distillation methods by which our Graph-KD performs best.

6 ACKNOWLEDGEMENTS

This work is supported by the National Science Foundation of China under Grant NSFC 61906194 and the National Key RD Program of China under Grant 25904.

REFERENCES

[1] Yan Bai, Yihang Lou, Feng Gao, Shiqi Wang, Yuwei Wu, and Ling-Yu Duan. 2018. Group-sensitive triplet embedding for vehicle reidentification. *IEEE Transactions on Multimedia* 20, 9 (2018), 2385–2399.

[2] Hao Chen, Benoit Lagadec, and Francois Bremond. 2019. Partition and Reunion: A Two-Branch Neural Network for Vehicle Re-identification.. In *CVPR Workshops*. 184–192.

[3] X. Chen, L. Zheng, C. Zhao, Q. Wang, and M. Li. 2020. RRGCCAN: Re-Ranking via Graph Convolution Channel Attention Network for Person Re-Identification. *IEEE Access* 8 (2020), 131352–131360. https://doi.org/10.1109/ACCESS.2020.3009653

[4] Yucheng Chen, Longlong Jing, Elahe Vahdani, Ling Zhang, Mingyi He, and Yingli Tian. 2019. Multi-camera Vehicle Tracking and Re-identification on AI City Challenge 2019.. In *CVPR Workshops*, Vol. 2.

[5] Ruihang Chu, Yifan Sun, Yadong Li, Zheng Liu, Chi Zhang, and Yichen Wei. 2019. Vehicle re-identification with viewpoint-aware metric learning. In *Proceedings of the IEEE International Conference on Computer Vision*. 8282–8291.

[6] Jian Du, Shanghang Zhang, Guanhang Wu, José MF Moura, and Soummya Kar. 2017. Topology adaptive graph convolutional networks. *arXiv preprint arXiv:1710.10370* (2017).

[7] Matthias Fey, Jan Eric Lenssen, Frank Weichert, and Heinrich Müller. 2018. Splinecnn: Fast geometric deep learning with continuous b-spline kernels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 869–877.

[8] Tommaso Furlanello, Zachary Lipton, Michael Tschannen, Laurent Itti, and Anima Anandkumar. 2018. Born again neural networks. In *International Conference on Machine Learning*. PMLR, 1607–1616.

[9] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. 2017. Neural message passing for quantum chemistry. In *International Conference on Machine Learning*. PMLR, 1263–1272.

[10] William L Hamilton, Rex Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *arXiv preprint arXiv:1706.02216* (2017).

[11] Bing He, Jia Li, Yifan Zhao, and Yonghong Tian. 2019. Part-regularized near-duplicate vehicle re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3997–4005.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[13] Lingxiao He, Xingyu Liao, Wu Liu, Xinchen Liu, Peng Cheng, and Tao Mei. 2020. FastReID: A Pytorch Toolbox for Real-world Person Re-identification. *arXiv preprint arXiv:2006.02631* (2020).

[14] Byeongho Heo, Minsik Lee, Sangdoo Yun, and Jin Young Choi. 2019. Knowledge transfer via distillation of activation boundaries formed by hidden neurons. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 3779–3787.

[15] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. 2015. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531* (2015).

[16] Deyi Ji, Haoran Wang, Hanzhe Hu, Weihao Gan, Wei Wu, and Junjie Yan. 2020. Context-Aware Graph Convolution Network for Target Re-identification. *arXiv preprint arXiv:2012.04298* (2020).

[17] Sameh Khamis, Cheng-Hao Kuo, Vivek K Singh, Vinay D Shet, and Larry S Davis. 2014. Joint learning for attribute-consistent person re-identification. In *European Conference on Computer Vision*. Springer, 134–146.

[18] Pirazh Khorramshahi, Amit Kumar, Neehar Peri, Sai Saketh Rambhatla, Jun-Cheng Chen, and Rama Chellappa. 2019. A dual-path model with adaptive attention for vehicle re-identification. In *Proceedings of the IEEE International Conference on Computer Vision*. 6132–6141.

[19] Pirazh Khorramshahi, Neehar Peri, Jun-Cheng Chen, and Rama Chellappa. 2020. The Devil Is in the Details: Self-supervised Attention for Vehicle Re-identification. In *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XIV (Lecture Notes in Computer Science)*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.), Vol. 12359. Springer, 369–386. https://doi.org/10.1007/978-3-030-58568-6_22

[20] Jangho Kim, SeongUK Park, and Nojun Kwak. 2018. Paraphrasing complex network: Network compression via factor transfer. *arXiv preprint arXiv:1802.04977* (2018).

[21] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).

[22] Shuang Li, Tong Xiao, Hongsheng Li, Bolei Zhou, Dayu Yue, and Xiaogang Wang. 2017. Person search with natural language description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1970–1979.

[23] Xuewei Li, Songyuan Li, Bourahla Omar, Fei Wu, and Xi Li. 2021. ResKD: Residual-Guided Knowledge Distillation. *IEEE Transactions on Image Processing* (2021).

[24] Yuqi Li, Yanghao Li, Hongfei Yan, and Jiaying Liu. 2017. Deep joint discriminative learning for vehicle re-identification and retrieval. In *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 395–399.

[25] Yutian Lin, Liang Zheng, Zhedong Zheng, Yu Wu, Zhilan Hu, Chenggang Yan, and Yi Yang. 2019. Improving person re-identification by attribute and identity learning. *Pattern Recognition* 95 (2019), 151–161.

[26] Hongye Liu, Yonghong Tian, Yaowei Yang, Lu Pang, and Tiejun Huang. 2016. Deep relative distance learning: Tell the difference between similar vehicles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2167–2175.

[27] Xinchen Liu, Wu Liu, Tao Mei, and Huadong Ma. 2017. Provid: Progressive and multimodal vehicle reidentification for large-scale urban surveillance. *IEEE Transactions on Multimedia* 20, 3 (2017), 645–658.

[28] Yihang Lou, Yan Bai, Jun Liu, Shiqi Wang, and Lingyu Duan. 2019. Veri-wild: A large dataset and a new method for vehicle re-identification in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3235–3243.

[29] Dechao Meng, Liang Li, Xuejing Liu, Yadong Li, Shijie Yang, Zheng-Jun Zha, Xingyu Gao, Shuhui Wang, and Qingming Huang. 2020. Parsing-based View-aware Embedding Network for Vehicle Re-Identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7103–7112.

[30] Wonpyo Park, Dongju Kim, Yan Lu, and Minsu Cho. 2019. Relational knowledge distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3967–3976.

[31] Baoyun Peng, Xiao Jin, Jiaheng Liu, Dongsheng Li, Yichao Wu, Yu Liu, Shunfeng Zhou, and Zhaoning Zhang. 2019. Correlation congruence for knowledge distillation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 5007–5016.

[32] Jingjing Qian, Wei Jiang, Hao Luo, and Hongyan Yu. 2020. Stripe-based and attribute-aware network: A two-branch deep model for vehicle re-identification. *Measurement Science and Technology* (2020).

[33] Meng Qu, Yoshua Bengio, and Jian Tang. 2019. Gmnn: Graph markov neural networks. In *International conference on machine learning*. PMLR, 5241–5250.

[34] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2008. The graph neural network model. *IEEE transactions on neural networks* 20, 1 (2008), 61–80.

[35] Yunsheng Shi, Zhengjie Huang, Shikun Feng, and Yu Sun. 2020. Masked label prediction: Unified massage passing model for semi-supervised classification. *arXiv preprint arXiv:2009.03509* (2020).

[36] Chi Su, Fan Yang, Shiliang Zhang, Qi Tian, Larry S Davis, and Wen Gao. 2015. Multi-task learning with low rank attribute embedding for person re-identification. In *Proceedings of the IEEE international conference on computer vision*. 3739–3747.

[37] Chi Su, Shiliang Zhang, Junliang Xing, Wen Gao, and Qi Tian. 2018. Multi-type attributes driven multi-camera person re-identification. *Pattern Recognition* 75 (2018), 77–89.

[38] Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, and Shengjin Wang. 2018. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). In *Proceedings of the European Conference on Computer Vision (ECCV)*. 480–496.

[39] Abhijit Suprem and Calton Pu. 2020. Looking GLAMORous: Vehicle Re-Id in Heterogeneous Cameras Networks with Global and Local Attention. *arXiv preprint arXiv:2002.02256* (2020).

[40] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).

[41] Guan'an Wang, Shuo Yang, Huanyu Liu, Zhicheng Wang, Yang Yang, Shuliang Wang, Gang Yu, Erjin Zhou, and Jian Sun. 2020. High-Order Information Matters: Learning Relation and Topology for Occluded Person Re-Identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

[42] Huibing Wang, Jinjia Peng, Dongyan Chen, Guangqi Jiang, Tongtong Zhao, and Xianping Fu. 2020. Attribute-guided Feature Learning Network for Vehicle Re-identification. *arXiv preprint arXiv:2001.03872* (2020).

[43] Zhongdao Wang, Liang Zheng, Yali Li, and Shengjin Wang. 2019. Linkage based face clustering via graph convolution network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1117–1125.

[44] Ancong Wu, Wei-Shi Zheng, and Jian-Huang Lai. 2019. Distilled camera-aware self training for semi-supervised person re-identification. *IEEE Access* 7 (2019), 156752–156763.

[45] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. 2020. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems* (2020).

[46] Cheng Yang, Jiawei Liu, and Chuan Shi. 2021. Extract the Knowledge of Graph Neural Networks and Go Beyond it: An Effective Knowledge Distillation Framework. *CoRR* abs/2103.02885 (2021). arXiv:2103.02885 https://arxiv.org/abs/2103.02885

[47] Yiding Yang, Jiayan Qiu, Mingli Song, Dacheng Tao, and Xinchao Wang. 2020. Distilling knowledge from graph convolutional networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7074–7083.

[48] Zhou Yin, Wei-Shi Zheng, Ancong Wu, Hong-Xing Yu, Hai Wan, Xiaowei Guo, Feiyue Huang, and Jianhuang Lai. 2018. Adversarial Attribute-Image Person Re-identification. In *Proceedings of the Twenty-Seventh International Joint Conference*

on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden, Jérôme Lang (Ed.). ijcai.org, 1100–1106. https://doi.org/10.24963/ijcai.2018/153

[49] Sergey Zagoruyko and Nikos Komodakis. 2017. Paying More Attention to Attention: Improving the Performance of Convolutional Neural Networks via Attention Transfer. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net. https://openreview.net/forum?id=Sks9_ajex

[50] Wentao Zhang, Xupeng Miao, Yingxia Shao, Jiawei Jiang, Lei Chen, Olivier Ruas, and Bin Cui. 2020. Reliable Data Distillation on Graph Convolutional Network. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. 1399–1414.

[51] Xinyu Zhang, Rufeng Zhang, Jiewei Cao, Dong Gong, Mingyu You, and Chunhua Shen. 2019. Part-guided attention learning for vehicle re-identification. *arXiv preprint arXiv:1909.06023* (2019).

[52] Zhizheng Zhang, Cuiling Lan, Wenjun Zeng, Xin Jin, and Zhibo Chen. 2020. Relation-Aware Global Attention for Person Re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 3186–3195.

[53] Yanzhu Zhao, Chunhua Shen, Huibing Wang, and Shengyong Chen. 2019. Structural analysis of attributes for vehicle re-identification and retrieval. *IEEE Transactions on Intelligent Transportation Systems* 21, 2 (2019), 723–734.

[54] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. 2020. Graph neural networks: A review of methods and applications. *AI Open* 1 (2020), 57–81.

[55] Jiahuan Zhou, Bing Su, and Ying Wu. 2020. Online Joint Multi-Metric Adaptation from Frequent Sharing-Subset Mining for Person Re-Identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2909–2918.

[56] Yi Zhou and Ling Shao. 2018. Viewpoint-Aware Attentive Multi-View Inference for Vehicle Re-Identification. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*. IEEE Computer Society, 6489–6498. https://doi.org/10.1109/CVPR.2018.00679

[57] Zaida Zhou, Chaoran Zhuge, Xinwei Guan, and Wen Liu. 2020. Channel Distillation: Channel-Wise Attention for Knowledge Distillation. *arXiv preprint arXiv:2006.01683* (2020).

[58] Jianqing Zhu, Huanqiang Zeng, Jingchang Huang, Shengcai Liao, Zhen Lei, Canhui Cai, and Lixin Zheng. 2019. Vehicle re-identification using quadruple directional deep learning features. *IEEE Transactions on Intelligent Transportation Systems* 21, 1 (2019), 410–420.

[59] Rixing Zhu, Jianwu Fang, Hongke Xu, Hongkai Yu, and Jianru Xue. 2020. DCDLearn: Multi-order Deep Cross-distance Learning for Vehicle Re-Identification. *arXiv preprint arXiv:2003.11315* (2020).