

Article

Not peer-reviewed version

---

# Unsupervised Vehicle Re-identification Method Based on Source-free Knowledge Transfer

---

Zhigang Song , [Daisong Li](#) , Zhongyou Chen , [Wengjin Yang](#) \*

Posted Date: 12 September 2023

doi: 10.20944/preprints202309.0754.v1

Keywords: vehicle re-identification; unsupervised domain adaptation; source-free knowledge transfer; pseudo-label; joint training



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Unsupervised Vehicle Re-Identification Method Based on Source-Free Knowledge Transfer

Zhigang Song<sup>1</sup>, Daisong Li<sup>1</sup>, Zhongyou Chen<sup>1</sup> and Wenqin Yang<sup>1,\*</sup>

<sup>1</sup> The Academy of Digital China (Fujian), Fuzhou University, Fuzhou 350108, China

\* Correspondence: yangwenqin@fzu.edu.cn; Tel.: 138 0508 4725

**Abstract:** The unsupervised domain-adaptive vehicle reidentification approach aims to transfer knowledge from a labeled source domain to an unlabeled target domain, however there are knowledge differences between the target domain and the source domain. To reduce domain differences, existing unsupervised domain-adaptive rerecognition methods generally require access to the source domain data to assist in re-training the target domain model. But for security reasons, in many cases, data between different domains cannot communicate. To this end, this paper proposes an unsupervised domain-adaptive vehicle re-identification method based on source-free knowledge transfer. First, by constructing a passive domain knowledge migration module, the target domain is consistent with the source domain model output to train a generator to generate the "source-like samples". Then, it can effectively reduce the model knowledge difference and improve the model generalization performance. According to the experiment and testing of VeRi776 and VehicleID, two mainstream public data sets in this field, the rank-k and mAP indexes obtained are both improved, and are suitable for object re-recognition tasks without source data.

**Keywords:** vehicle re-identification; unsupervised domain adaptation; source-free knowledge transfer; pseudo-label; joint training

## 1. Introduction

Vehicle re-identification (Re-ID) refers to judging whether vehicle images captured in non-overlapping areas belong to the same vehicle in a traffic monitoring scene within a specific range. Recently, vehicle re-identification methods based on supervised learning have made great progress [1–4]. However, the supervised learning method mainly has the following problems: 1) Extremely dependent on complete labels, that is, the labels of training data from multiple non-overlapping cameras, annotating all large-scale unlabeled data, which is time-consuming and labor-intensive; 2) These methods perform well in the original task (source domain), but when deployed in a new environment (target domain), the performance will drop significantly due to the presence of domain bias.

To overcome these problems, researchers began to focus on the research of vehicle re-identification methods based on unsupervised domain-adaptive [5–10], that is, trying to transfer images from a well-marked source domain datasets to an unlabeled target domain datasets through knowledge transfer. Isola et al. [11] used the generated images to train the ReID model earlier, by preserving the identity information from the well-labeled domain, while learning the style of the unlabeled domain to improve the performance of the model in the unlabeled target domain. Peng et al. [12] proposed a Progressive Adaptation Learning algorithm for vehicle re-identification. This method utilizes the source domain to generate "pseudo-target samples" through the Generative Adversarial Network (GAN). It employs a dynamic sampling strategy during training to mitigate domain discrepancies. Zheng et al. [13] proposed a viewpoint-aware clustering algorithm. It leverages a pre-trained vehicle orientation predictor to predict the orientations of vehicles, assign directional pseudo-labels, and first clusters vehicles with the same perspective. Subsequently, it clusters vehicles with different perspectives, thereby enhancing the performance of the vehicle re-identification model. Wang et al. [14] proposed a progressive learning method named PLM for vehicle re-identification in unknown

domains. The method utilizes domain adaptation and multi-scale attention network to smooth domain bias, trains a reID model, and introduces a weighted label smoothing loss to improve performance. The above methods preserve the identity information from a well-labeled source domain while learning the style of the unlabeled target domain, but usually face the problem that when learning an adaptive model on the target domain, it is unavoidable to visit the source domain to generate new samples for subsequent fine-tuning. However, due to data ownership and privacy issues in different domains, inter-domain data cannot be communicated in many cases, and the target domain model cannot directly access the source domain data, and the adaptive performance of the model will be greatly affected.

To this end, this paper proposes an unsupervised domain-adaptive vehicle re-identification method based on source-free knowledge transfer, that is, a target domain sample is given, and the migration image of the sample is obtained through the generator, and the two images are respectively provided to the target model and the source model, let the difference between the image pair compensate for the knowledge difference between the domain models, so that the output of the two domain models is similar, and then train the generator by constraining the output similarity, and the target domain data can be transformed to have the style of the source domain. Therefore, this "source-like samples" can replace the role played by the source domain data in the model adaptation of the target domain, and since the generated sample content is provided by the target domain, it is more affinity in the process of model adaptation, which helps solve the problem that the target domain cannot access the data of the source domain. The method can be divided into two stages:

(1) In the first stage, a source-free knowledge transfer module is built, which only uses the source domain model and the target domain model trained by unlabeled target domain data as supervision, and trains a generator to generate "source-like samples" without accessing the source domain data. "source-like samples" match the style of the source domain and the content of the target domain.

(2) In the second stage, a progressive joint training strategy is adopted to gradually learn the adaptive model by inputting different proportions of "source-like samples" and target domain data. This process can be regarded as a means of data expansion. Compared with the target domain directly acting on the source domain model, the "source-like samples" containing source domain knowledge can be more compatible with the model, and the domain can be effectively reduced in iterative training. In order to achieve the purpose of improving the generalization performance of the model.

The contributions of this paper can be summarized as follows:

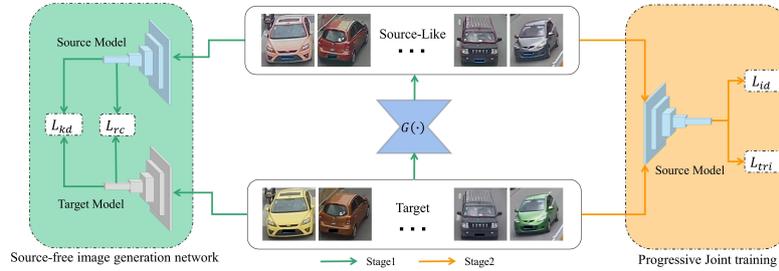
(1) We propose an unsupervised domain-adaptive vehicle re-identification method based on source-free knowledge transfer. It does not need to access the source domain data, and uses the hidden domain difference information in the source domain model and the target domain model to constrain the generator to generate "source-like samples". Samples are used as a data augmentation method for vehicle re-identification tasks to assist model training.

(2) We propose a progressive joint training strategy of "source-like samples" and the target domain. "source-like samples" adapt to the source domain model in the same style, and match the target domain data with the same content, as the source domain model and target domain data. An intermediate hub for adaptation, which mitigates domain differences and thus improves model performance.

## 2. Method

In this section, we give a detailed description of the proposed method. The schematic diagram of the method is shown in Figure 1, Only need to be given a source model and a target model, through the source-free knowledge transfer module constructed in this paper, a generator  $G(\cdot)$  can be trained to generate "source-like samples"  $\tilde{x}$ . This sample contains source domain knowledge, which is more affinity than the target domain directly acting on the source domain model. It can be used as data expansion to assist in training the target model, which helps to improve model performance. Then, a progressive joint training strategy is further adopted to use this sample and target domain data for training, and by controlling the proportion of "source-like samples" and the original sample in the

target domain, the model performance may be degraded due to the high proportion of noise samples. Because the method in this paper does not need to access the source domain data, it can overcome the limitation that the existing unsupervised domain-adaptive methods need to access the source domain, and avoid the security and transmission problems that may be caused by accessing the source domain data.



**Figure 1.** Schematic diagram of unsupervised domain-adaptive vehicle re-identification method based on source-free knowledge transfer

### 2.1. Pretrained source and target models

Among the existing unsupervised domain-adaptive methods[15], the usual practice is to train the model  $F(\cdot | \theta)$  on the source domain first, where  $\theta$  represents the parameters of the current model, and then transfer the model to the target domain for learning.

The method in this paper does not need to access the source data, and its source domain model is expressed as:

$$\mathbb{D}_s = \{x_i^s, y_i^s\} |_{i=1}^{N_s} \quad (1)$$

The target domain model is expressed as:

$$\mathbb{D}_t = \{x_i^t\} |_{i=1}^{N_t} \quad (2)$$

The experimental steps are as follows:

First, access the source domain to train the source model as well as a learnable source domain classifier  $C^s : f^s \rightarrow \{1, 2, \dots, M_s\}$ , where  $M_s$  represents the number of sample identities.

Second, optimize using the identity classification loss  $\mathcal{L}_{id}^s(\theta)$  and triplet loss  $\mathcal{L}_{tri}^s(\theta)$ [16] composed of the cross-entropy loss function  $L_{ce}$ , as shown in equations (3) and (4):

$$\mathcal{L}_{id}^s(\theta) = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathcal{L}_{ce}(C^s(F^s(x_i^s | \theta)), y_i^s) \quad (3)$$

$$\mathcal{L}_{tri}^s(\theta) = \frac{1}{N_s} \sum_{i=1}^{N_s} \max(0, \|F^s(x_i^s | \theta) - F^s(x_{(i,p)}^s | \theta)\| + m - \|F^s(x_i^s | \theta) - F^s(x_{(i,n)}^s | \theta)\|) \quad (4)$$

where  $\|\cdot\|$  represents the  $L_2$ -norm; the subscripts  $(i, p)$  and  $(i, n)$  represents the positive sample and the negative sample of the  $i$ -th sample, respectively; and  $m = 0.5$  represents the distance margin of triplet loss. The overall loss  $\mathcal{L}^s(\theta)$  is therefore calculated as:

$$\mathcal{L}^s(\theta) = \mathcal{L}_{id}^s(\theta) + \lambda^s \mathcal{L}_{tri}^s(\theta) \quad (5)$$

where  $\lambda^s$  represents the weight of two losses. After obtaining the source model, we learn a target model by loading the source model parameters, clustering the target domain, and then predicting the pseudo labels  $\hat{y}_i^t$ . The overall loss  $\mathcal{L}^t(\theta)$  is therefore calculated as:

$$\mathcal{L}^t(\theta) = \mathcal{L}_{id}^t(\theta) + \lambda^t \mathcal{L}_{tri}^t(\theta) \quad (6)$$

## 2.2. Source-free image generation module

The existing unsupervised domain adaptive vehicle re-identification methods usually need to access the source domain data, and transfer the well-marked source domain data to the unmarked target domain style through style-based transfer[17,18] or generative confrontation network[19,20], so as to smooth domain bias to better apply source domain models to target domain data. However, when some data has security and privacy restrictions, it is extremely difficult to access the source domain data. In order to solve this problem, this paper constructs a source-free image generation module, which aims to force the generator to generate "source-like samples" from the target domain data with the style of the source domain through the implicit domain information constraints in the model, generated to bridge the knowledge gap between models.

As shown in Figure 2, the first line is the target samples image, and the second line is the "source-like samples" obtained by the target image through the source-free image generation module. The biggest feature of this sample is that its content matches the target domain data, and its style can match the source domain model. It can act as a bridge between source and target domains. In the subsequent model optimization process, the joint training of this sample and the target domain sample can effectively improve the generalization performance of the vehicle re-identification model.



**Figure 2.** Schematic diagram of comparison between source-like samples and target samples.

The source-free image generation module aims to train the image generator  $G(\cdot)$  to generate "source-like samples" with the style of the source domain through the domain information implicit in the model. These samples replace the source domain data to complete the matching source model. To describe the knowledge adapted in "source-like samples", in addition to the traditional knowledge distillation loss  $\mathcal{L}_{kd}$  a channel-level relational consistency loss  $\mathcal{L}_{rc}$  is also introduced. It can focus on the relative channel relationship between feature maps of the target domain and "source-like samples". Therefore, the total loss  $\mathcal{L}_{SFIG}$  is shown in equation (7):

$$\mathcal{L}_{SFIG} = \mathcal{L}_{kd} + \mathcal{L}_{rc} \quad (7)$$

In the following sections, the details of the two losses will be introduced in detail.

### 2.2.1. knowledge distillation loss

In our proposed source-free image generation network, we utilize a combination  $F_s(G(x_i^t))$  of the source model and generator to describe the knowledge adapted in the target model. This approach can be considered as a special application of knowledge distillation. Our aim is to extract knowledge differences between two domains into the generator. In this case, we compose the knowledge

distillation loss  $\mathcal{L}_{kd}$  by the output  $\rho(F^s(\tilde{x}))$  obtained by feeding "source-like samples" into the source model and the output  $\rho(F^t(x_i^t))$  obtained by feeding target domain samples into the target model.

$$\mathcal{L}_{kd} = \mathcal{D}_{kl}(\rho(F^s(\tilde{x})), \rho(F^t(x_i^t))) \quad (8)$$

where  $\mathcal{D}_{kl}(\cdot)$  represents the Kullback–Leibler divergence(KL).

### 2.2.2. Channel-level relational consistency loss

In unsupervised domain-adaptive tasks, it is usually assumed that there is a fixed classifier, so it can be considered that the global features obtained by the target domain through the target model should be similar to the global features obtained by "source-like samples" through the source model. To promote similar channel-level relationships between feature maps  $f^s \rightarrow F^s(\tilde{x})$  and  $f^t \rightarrow F^t(x_i^t)$ , a relation consistency loss is used to constrain.

Previous knowledge distillation work is usually constrained by maintaining batch-level or pixel-level relationships[21,22]. However, this constrained approach is not suitable for the current task. First of all, the batch-level relationship cannot well supervise the generation task of each image, which will cause damage to the generated effect. Second, the effectiveness of pixel-level relationships will be greatly reduced after global pooling. Compared with the two, the channel-level relationship[23] is computed on a per-image basis and is not affected by global pooling. Therefore, the channel-level relationship is more suitable for computing  $\mathcal{L}_{rc}$ .

Given the feature map  $f^s$  of "source-like samples" and the feature map  $f^t$  of target domain, we resize them into feature vectors  $\mathcal{F}^s$  and  $\mathcal{F}^t$ , as shown in equation (9) and equation (10):

$$f^s \in \mathbb{R}^{D*H*W} \rightarrow \mathcal{F}^s \in \mathbb{R}^{D*HW} \quad (9)$$

$$f^t \in \mathbb{R}^{D*H*W} \rightarrow \mathcal{F}^t \in \mathbb{R}^{D*HW} \quad (10)$$

where  $D, H$ , and  $W$ , represent the feature map depth(number of channels), height, and width, respectively. Next, we compute their channel-level self correlation, the gram matrix, as shown in equation (11):

$$\mathcal{G}^s = \mathcal{F}^s \cdot (\mathcal{F}^s)^T, \mathcal{G}^t = \mathcal{F}^t \cdot (\mathcal{F}^t)^T \quad (11)$$

where  $\mathcal{G}^s, \mathcal{G}^t \in \mathbb{R}^{D*D}$ . Like other similarity-preserving losses for knowledge distillation, we apply row-level  $L_2$ -norm, as shown in equation (12):

$$\tilde{\mathcal{G}}_{[i,:]}^s = \frac{\mathcal{G}_{[i,:]}^s}{\|\mathcal{G}_{[i,:]}^s\|_2}, \tilde{\mathcal{G}}_{[i,:]}^t = \frac{\mathcal{G}_{[i,:]}^t}{\|\mathcal{G}_{[i,:]}^t\|_2} \quad (12)$$

where  $[i, :]$  represents row  $i$  in the matrix. Finally, the channel-level relational consistency loss  $\mathcal{L}_{rc}$  is the mean squared error(MSE) between the normalized Gram matrices, as shown in equation (13):

$$\mathcal{L}_{rc} = \frac{1}{D} \|\mathcal{G}^s - \mathcal{G}^t\|_F^2 \quad (13)$$

### 2.3. Progressive joint training strategy

In the previous section, the target domain data was used to generate "source-like samples" through the source-free knowledge transfer module. Since this sample combines the style of the source domain and the vehicle content of the target domain, it can be used for training to improve the target. The performance degradation caused by deploying domain data to the source domain model improves the generalization performance of the model. In addition, during the process of model adaptation, due to issues such as image style and image quality, "source-like samples" may be regarded as noise by the model. Therefore, a progressive joint training strategy is introduced to control the feeding ratio

of "source-like samples" and the target domain, which can effectively prevent the one-time input of "source-like samples" from causing too much noise and destroying the model performance. In addition, as the proportion of "source-like samples" increases, the adaptability of the model to the fusion of two samples is also enhanced, so that more discriminative features are learned on the target domain. The maximum ratio of "source-like samples" to the target domain is 1:1.

During the training process, the "source-like samples" and target domain data are processed by using the pre-trained source model of vehicle re-identification with good performance to output high-dimensional features. Most of the previous methods choose K-Means to generate clusters, which need to be initialized by cluster centroids. However, it is uncertain how many categories are required in the target domain. Therefore, DBSCAN[24] is chosen as the clustering method. Specifically, instead of using a fixed clustering radius, this paper adopts a dynamic clustering radius calculated by K-Nearest Neighbors (KNN). After DBSCAN, in order to filter noise, some of the most reliable samples are selected for soft label assignment according to the distance between sample features and cluster centroids. For the proposed method, samples satisfying  $\|f_i, c_{f_i}\| < \gamma$  are used for the next iteration, where  $f_i$  is the feature of the  $i$ -th image,  $c_{f_i}$  is the feature of the centroid of the cluster to which  $f_i$  belongs, and  $\gamma$  represents the metric radius belonging to the same category.

### 3. Experiment

#### 3.1. Experimental environment settings

Currently commonly used open source deep learning frameworks include Caffe, Tensorflow, pytorch, etc. The method proposed in this paper is trained in the pytorch framework, including the ablation experiment is also conducted in pytorch. Compared with other frameworks, python, the development language of Pytorch, has the biggest advantage of supporting dynamic neural networks, and the framework is more intuitive and concise as a whole. All models are trained on the RTX3060 graphics card, the initial learning rate is set to 0.0005, and the weight decay is set to 0.0005. The total number of iterations is set to 50 and optimized using the Adam optimizer. In order to facilitate the experiment, the model training framework proposed in [5] is used here to obtain the initial source domain model and target domain model. The source domain data is accessed here only to provide the initial source domain model, and the source domain data will no longer be accessed in subsequent experiments. The main architecture of the model adopts ResNet50[25]. Due to memory issues, the generator in the source-free knowledge transfer module was built using a modified CycleGAN architecture with 3 residual blocks. The source and target domain models are fixed during training. During the training process, "source-like samples" generated by the source-free knowledge transfer module are used for training in various ratios with target domain data. The ratios between "source-like samples" and target domain data are set as 1:5, 1:4, 1:3, 1:2, and 1:1. The clustering method employed is DBSCAN.

#### 3.2. Datasets setting and evaluation index level

All experiments are conducted on two datasets, VeRi776[26] and VehicleID[27]. For each vehicle in the test datasets, a query image is selected from the perspective of each camera where it is located, and the input images in the experiment are modified to 240×240, and the number of input images in each batch is 32. Training The input images are preprocessed by random horizontal flipping and random erasing. The evaluation indicators used in the experiment are Rank-1 and Rank-5 in the cumulative matching features, as well as the mean Average Precision (mAP).

#### 3.3. Experimental results and analysis

This subsection compares the proposed method with existing unsupervised methods. The tables in Tables 1–3 show the comparison results. In Table 1, the initial source domain model is trained on

the VehicleID datasets, and the target domain is VeRi776. In Tables 2 and 3, the initial source domain model is trained on the VeRi776 datasets, and the target domain is VehicleID.

### 3.3.1. Experimental results and analysis on VehicleID→VeRi776

The experimental results of this method on VehicleID→VeRi776 are shown in Table 1. Obviously, the proposed method has achieved good experimental results on the VeRi776 datasets, with Rank-1, Rank-5, and mAP being 74.4%, 82.1%, and 37.9%, respectively. In particular, PUL[28], HHL[29] and other methods are based on human-based unsupervised re-identification methods. Since most of the current vehicle re-identification is based on supervised methods, they can be used for unsupervised vehicle re-identification for comparison. Methods are scarce, so these pedestrian-based unsupervised re-ID methods are applied to vehicle re-ID to compare with the proposed method.

**Table 1.** Comparative experimental results of the method in this paper on VeRi776

Methods	VeRi776		
	Rank-1(%)	Rank-5(%)	mAP(%)
PUL[28]	55.24	66.27	17.06
SPGAN[30]	57.4	70.0	16.4
HHL[29]	56.20	67.61	17.52
ECN[31]	60.8	70.9	27.7
UDAP[5]	73.9	81.5	35.8
PAL[12]	68.17	79.91	<b>42.04</b>
Direct Transfer	62.1	73.9	27.6
ours	<b>74.4</b>	<b>82.1</b>	37.9

It is worth mentioning that the unsupervised vehicle re-identification method PAL[12] was also found through research. It is a method of learning cross-domain multi-semantic knowledge for unsupervised vehicle re-identification. Its architecture is quite similar to the method proposed in this paper. In comparison, our proposed method achieves an improvement of 6.3% in Rank-1 and 2.2% in Rank-5, with a slightly lower mAP compared to PAL. Experiments prove that, compared with the PAL that still needs to access the source domain during the migration process, the proposed source-free knowledge transfer module can generate reliable "Source-like samples" without accessing the source domain data, and through joint training The strategy effectively smoothes the domain difference between the target domain and the source domain to gain performance improvement.

### 3.3.2. Experimental results and analysis on VeRi776 → VehicleID

The experimental results of this method on VeRi776 → VehicleID are shown in Tables 2 and 3. It can be seen that the proposed method has also achieved the best experimental results on the VehicleID datasets. When the test size is 800, 1600, 2400, and 3200, the Rank-1 are 52.76%, 47.65%, 43.87%, and 41.77%, respectively, the Rank-5 are 67.29%, 63.83%, 62.43%, 60.42%, mAP are 58.33%, 53.72%, 50.42%, 47.29%. It is worth noting that the method proposed in this paper has a huge improvement compared with the direct application of the target domain on the source model (direct transfer). Taking test size=800 as an example, Rank-1 and mAP are increased by 13.2% and 15.32%, respectively. In addition, since the scale of VeRi776 is much smaller than that of VehicleID, there are very few generalization studies on the realization of vehicle re-identification tasks from small data sets to large data sets, so we can only select some from a small number of research works. Relatively prominent methods are compared with this paper. It is not difficult to see from Tables 2 and 3 that compared with PAL, which is most similar to the method in this paper, it shows a stable improvement in different test sizes.

**Table 2.** Comparative experimental results of the method in this paper on VehicleID (part 1)

Methods	Test size = 800			Test size = 1600		
	Rank-1(%)	Rank-5(%)	mAP(%)	Rank-1(%)	Rank-5(%)	mAP(%)
PUL[28]	40.03	46.03	43.9	33.83	49.72	37.68
CycleGAN[19]	37.29	58.56	42.32	30.00	49.96	34.92
PAL[12]	50.25	64.91	53.50	44.25	60.95	48.05
Direct Transfer	39.56	56.03	43.01	35.01	50.84	39.17
ours	52.76	67.29	58.33	47.65	63.83	53.72

**Table 3.** Comparative experimental results of the method in this paper on VehicleID (part 2)

Methods	Test size = 2400			Test size = 3200		
	Rank-1(%)	Rank-5(%)	mAP(%)	Rank-1(%)	Rank-5(%)	mAP(%)
PUL[28]	30.90	47.18	34.71	28.86	43.41	32.44
CycleGAN[19]	27.15	46.52	31.86	24.83	42.17	29.17
PAL[12]	41.08	59.12	45.14	38.19	55.32	42.13
Direct Transfer	31.05	48.52	34.72	28.12	42.98	31.99
ours	43.87	62.43	50.42	41.77	60.42	47.29

Through VeRi776→VehicleID, the experiments on VehicleID→VeRi776 have achieved outstanding performance, verifying the effectiveness of the proposed method, indicating that the method can also help the unsupervised domain- adaptive vehicle re-identification method under the setting of source-free domain improve performance, but compared with some supervised learning methods, there is still a certain gap, and continuous exploration is still needed.

### 3.4. Ablation experiment

In this section, in order to verify the effectiveness of each module of the proposed method, an ablation experiment was conducted on the VeRi776 datasets.

#### 3.4.1. Validation of the loss function of the source-free image generation module

In the source-free knowledge transfer module, the generator is constrained by the distillation loss and the channel-level relational consistency loss, forcing the target domain image to generate "source-like samples" with the style of the source domain under the guidance of the source domain model and the target domain model. The impact of the loss function of the source-free domain image generation module on performance is shown in Table 4. It can be clearly seen that when the two losses work together, the generated "source-like samples" are the best for training as image augmentation. Compared with  $\mathcal{L}_{rc}$  alone, Rank-1 increased by 4.2%, mAP increased by 3.5%; compared with  $\mathcal{L}_{kd}$  alone, Rank-1 increased by 6.8%, mAP increased by 5.8%. In addition, the effect of using only  $\mathcal{L}_{rc}$  is also slightly higher than using only  $\mathcal{L}_{kd}$ . This is due to the fact that focusing on relative channel relationships can better preserve foreground objects (less blurry and more prominent) while transferring the overall image style, resulting in higher recognition accuracy.

**Table 4.** The loss function of the source-free image generation module's impact on performance.

Loss Function	VeRi776		
	Rank-1(%)	Rank-5(%)	mAP(%)
$\mathcal{L}_{kd}$	67.6	78.6	32.1
$\mathcal{L}_{rc}$	70.2	80.3	34.4
$\mathcal{L}_{kd} + \mathcal{L}_{rc}$	74.4	82.1	37.9

### 3.4.2. Validation of the effectiveness of "source-like samples"

The impact of "source-like samples" on performance is shown in Table 5. Through data comparison, it is not difficult to see that when the target domain data is directly deployed to the source domain model in "direct migration", the performance deteriorates due to domain differences. Rank-1 dropped by 29.3%, mAP dropped by 42.9%. After only using "source-like samples" for training, the performance of the model has been significantly improved, with a 10.4% increase in Rank-1 and a 7.5% increase in mAP. It is worth mentioning that after the joint training of "source-like samples" and target domain data, the model has been further improved. This fully demonstrates that using the characteristics of "source-like samples" to match the content of the target domain and the style to match the source domain can help smooth inter-domain deviations during training, thereby improving the generalization performance of the model.

**Table 5.** The impact of "source-like samples" on performance.

Type	VeRi776		
	Rank-1(%)	Rank-5(%)	mAP(%)
Supervised Learning	91.4	96.2	70.5
Direct Transfer	62.1	73.9	27.6
Source-Like Samples	72.5	81.5	35.1
Joint Training	74.4	82.1	37.9

### 3.4.3. Validation of progressive joint training strategy

The performance impact results of the progressive joint training strategy are shown in Table 6, which are studied by inputting different ratios of "source-like samples" to the target domain data. Comparing the data, it can be found that when the initial proportion of "source-like samples" is small, the performance improvement of the model is also very limited. Rank-1 has increased by 5.3%, and mAP has increased by 3.7%. After the input ratio is 1:3, the performance change of the model tends to be flat, and it reaches the best when the input ratio is 1:2, the Rank-1 is increased by 12.3%, and the mAP is increased by 9.7%. This shows that the joint training strategy can gradually alleviate the domain differences in continuous training iterations, thereby improving the model performance.

**Table 6.** The Impact of progressive joint training strategy on performance.

Feed Ratio	VeRi776		
	Rank-1(%)	Rank-5(%)	mAP(%)
1:5	67.4	76.5	31.3
1:4	69.5	78.9	33.2
1:3	72.6	80.9	36.1
1:2	74.4	82.1	37.9
1:1	73.6	81.7	37.4

## 4. Conclusion

In this paper, a data augmentation method based on source-free knowledge transfer is proposed to assist unsupervised vehicle re-identification. "Source-like samples" are generated through a source-free knowledge transfer module, and target domain adaptation is accomplished using a joint training strategy. The proposed method is comprehensively tested on benchmark datasets VeRi776 and VehicleID, respectively. Compared with other existing unsupervised vehicle re-identification methods, the proposed method has better performance. Ablation experiments also demonstrate the effectiveness of the proposed individual components. In the future, we will continue to explore the cross-domain problem in unsupervised vehicle re-identification.

**Author Contributions:** Conceptualization, Z.S. and D.L.; methodology, Z.S., D.L. and Z.C.; software, D.L. and Z.C.; validation, Z.S., D.L. and Z.C.; formal analysis, Z.S.; investigation, Z.S.; resources, Z.S.; data curation, D.L. and Z.C.; writing—original draft preparation, D.L. and Z.C.; writing—review and editing, Z.S.; visualization, W.Y.; supervision, Z.S. and W.Y.; project administration, Z.S. and W.Y.; funding acquisition, W.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Fujian Science and Technology Plan: The special project of central government guiding local science and technology development(No.2022L3003)

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Huang, L.; Yang, Q.; Wu, J.; Huang, Y.; Wu, Q.; Xu, J. Generated data with sparse regularized multi-pseudo label for person re-identification. *IEEE Signal Processing Letters* **2020**, *27*, 391–395.
2. Bai, Y.; Lou, Y.; Gao, F.; Wang, S.; Wu, Y.; Duan, L.Y. Group-sensitive triplet embedding for vehicle reidentification. *IEEE Transactions on Multimedia* **2018**, *20*, 2385–2399.
3. Zhao, Y.; Shen, C.; Wang, H.; Chen, S. Structural analysis of attributes for vehicle re-identification and retrieval. *IEEE Transactions on Intelligent Transportation Systems* **2019**, *21*, 723–734.
4. Chen, Y.; Guo, K.; Liu, F.; Huang, Y.; Qi, Z. Supervised contrastive vehicle quantization for efficient vehicle retrieval. In Proceedings of the Proceedings of the 2022 International Conference on Multimedia Retrieval, 2022, pp. 44–48.
5. Song, L.; Wang, C.; Zhang, L.; Du, B.; Zhang, Q.; Huang, C.; Wang, X. Unsupervised domain adaptive re-identification: Theory and practice. *Pattern Recognition* **2020**, *102*, 107173.
6. Yi, Z.; Zhang, H.; Tan, P.; Gong, M. Dualgan: Unsupervised dual learning for image-to-image translation. In Proceedings of the Proceedings of the IEEE international conference on computer vision, 2017, pp. 2849–2857.
7. Yan, T.; Guo, H.; Liu, S.; Zhao, C.; Tang, M.; Wang, J. Unsupervised domain adaptive re-identification with feature adversarial learning and self-similarity clustering. In Proceedings of the Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event, January 10–15, 2021, Proceedings, Part IV. Springer, 2021, pp. 20–35.
8. Wei, R.; Gu, J.; He, S.; Jiang, W. Transformer-Based Domain-Specific Representation for Unsupervised Domain Adaptive Vehicle Re-Identification. *IEEE Transactions on Intelligent Transportation Systems* **2022**, *24*, 2935–2946.
9. Zhu, W.; Peng, B. Manifold-based aggregation clustering for unsupervised vehicle re-identification. *Knowledge-Based Systems* **2022**, *235*, 107624.
10. Wang, Y.; Wei, Y.; Ma, R.; Wang, L.; Wang, C. Unsupervised vehicle re-identification based on mixed sample contrastive learning. *Signal, Image and Video Processing* **2022**, *16*, 2083–2091.
11. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.
12. Peng, J.; Wang, Y.; Wang, H.; Zhang, Z.; Fu, X.; Wang, M. Unsupervised vehicle re-identification with progressive adaptation. *arXiv preprint arXiv:2006.11486* **2020**.
13. Zheng, A.; Sun, X.; Li, C.; Tang, J. Aware progressive clustering for unsupervised vehicle re-identification. *IEEE Transactions on Intelligent Transportation Systems* **2021**, *23*, 11422–11435.
14. Wang, Y.; Peng, J.; Wang, H.; Wang, M. Progressive learning with multi-scale attention network for cross-domain vehicle re-identification. *Science China Information Sciences* **2022**, *65*, 160103.
15. Ge, Y.; Chen, D.; Li, H. Mutual mean-teaching: Pseudo label refinery for unsupervised domain adaptation on person re-identification. *arXiv preprint arXiv:2001.01526* **2020**.
16. Hermans, A.; Beyer, L.; Leibe, B. In defense of the triplet loss for person re-identification. *arXiv preprint arXiv:1703.07737* **2017**.
17. Gatys, L.A.; Ecker, A.S.; Bethge, M. Image style transfer using convolutional neural networks. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2414–2423.

18. Huang, X.; Belongie, S. Arbitrary style transfer in real-time with adaptive instance normalization. In Proceedings of the Proceedings of the IEEE international conference on computer vision, 2017, pp. 1501–1510.
19. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the Proceedings of the IEEE international conference on computer vision, 2017, pp. 2223–2232.
20. Hoffman, J.; Tzeng, E.; Park, T.; Zhu, J.Y.; Isola, P.; Saenko, K.; Efros, A.; Darrell, T. Cycada: Cycle-consistent adversarial domain adaptation. In Proceedings of the International conference on machine learning. Pmlr, 2018, pp. 1989–1998.
21. Tung, F.; Mori, G. Similarity-preserving knowledge distillation. In Proceedings of the Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 1365–1374.
22. Li, Z.; Jiang, R.; Aarabi, P. Semantic relation preserving knowledge distillation for image-to-image translation. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVI 16. Springer, 2020, pp. 648–663.
23. Hou, Y.; Zheng, L. Visualizing adapted knowledge in domain transfer. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 13824–13833.
24. Khan, K.; Rehman, S.U.; Aziz, K.; Fong, S.; Sarasvady, S. DBSCAN: Past, present and future. In Proceedings of the The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014). IEEE, 2014, pp. 232–238.
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
26. Liu, X.; Liu, W.; Mei, T.; Ma, H. Provid: Progressive and multimodal vehicle reidentification for large-scale urban surveillance. *IEEE Transactions on Multimedia* **2017**, *20*, 645–658.
27. Liu, H.; Tian, Y.; Yang, Y.; Pang, L.; Huang, T. Deep relative distance learning: Tell the difference between similar vehicles. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2167–2175.
28. Fan, H.; Zheng, L.; Yan, C.; Yang, Y. Unsupervised person re-identification: Clustering and fine-tuning. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **2018**, *14*, 1–18.
29. Zhong, Z.; Zheng, L.; Li, S.; Yang, Y. Generalizing a person retrieval model hetero-and homogeneously. In Proceedings of the Proceedings of the European conference on computer vision (ECCV), 2018, pp. 172–188.
30. Deng, W.; Zheng, L.; Ye, Q.; Kang, G.; Yang, Y.; Jiao, J. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 994–1003.
31. Zhong, Z.; Zheng, L.; Luo, Z.; Li, S.; Yang, Y. Invariance matters: Exemplar memory for domain adaptive person re-identification. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 598–607.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.