

## Article

# Efficient and Deep Vehicle Re-Identification Using Multi-Level Feature Extraction

Zakria <sup>1,\*</sup>, Jingye Cai <sup>1</sup>, Jianhua Deng <sup>1</sup>, Muhammad Umar Aftab <sup>1</sup> ,  
Muhammad Saddam Khokhar <sup>2</sup>  and Rajesh Kumar <sup>3</sup> 

<sup>1</sup> School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 610054, China; jycail@uestc.edu.cn (J.C.); jianhua.deng@uestc.edu.cn (J.D.); ms.umaraftab@yahoo.com (M.U.A.)

<sup>2</sup> School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, Jiangsu 212013, China; saddam\_khokhar@hotmail.com

<sup>3</sup> School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 610054, China; rajakumarlohano@gmail.com

\* Correspondence: zakria.uestc@hotmail.com

Received: 17 February 2019; Accepted: 21 March 2019; Published: 27 March 2019



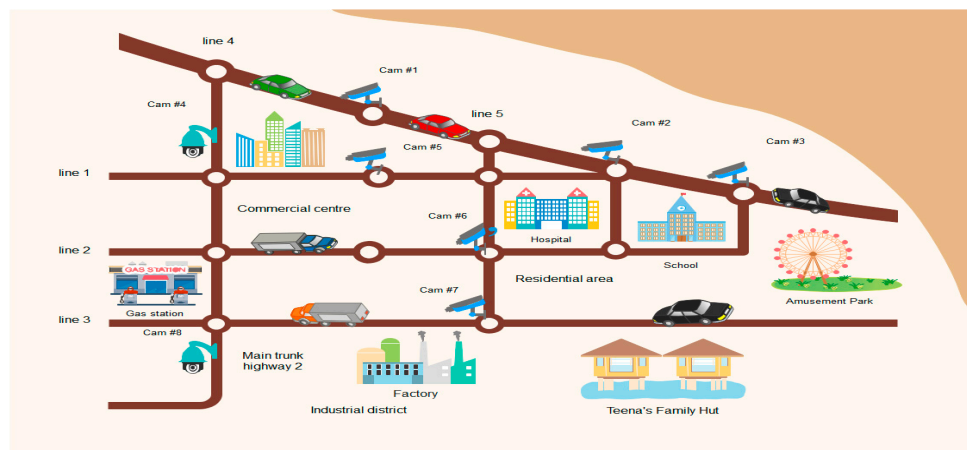
**Abstract:** The intelligent transportation system is currently an active research area, and vehicle re-identification (Re-Id) is a fundamental task to implement it. It determines whether the given vehicle image obtained from one camera has already appeared over a camera network or not. There are many possible practical applications where the vehicle Re-Id system can be employed, such as intelligent vehicle parking, suspicious vehicle tracking, vehicle incident detection, vehicle counting, and automatic toll collection. This task becomes more challenging because of intra-class similarity, viewpoint changes, and inconsistent environmental conditions. In this paper, we propose a novel approach that re-identifies a vehicle in two steps: first we shortlist the vehicle from a gallery set on the basis of appearance, and then in the second step we verify the shortlisted vehicle's license plates with a query image to identify the targeted vehicle. In our model, the global channel extracts the feature vector from the whole vehicle image, and the local region channel extracts more discriminative and salient features from different regions. In addition to this, we jointly incorporate attributes like model, type, and color, etc. Lastly, we use a siamese neural network to verify license plates to reach the exact vehicle. Extensive experimental results on the benchmark dataset VeRi-776 demonstrate the effectiveness of the proposed model as compared to various state-of-the-art methods.

**Keywords:** vehicle re-identification; overlapping and non-overlapping cameras; global and local features; deep learning; license plate

## 1. Introduction

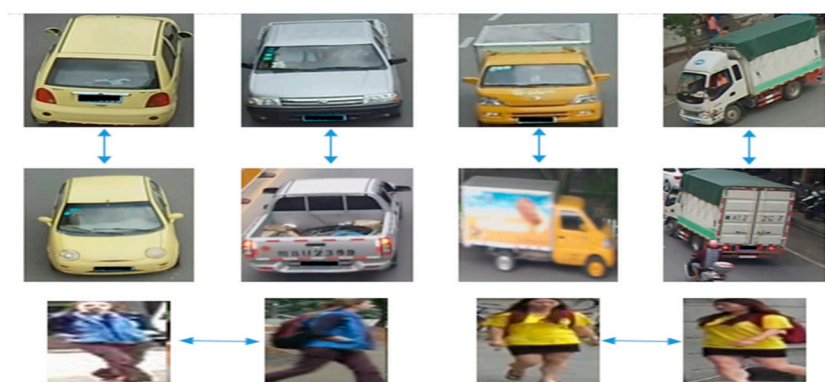
Nowadays, the security and safety of objects and public place is an important issue in society. Government and private sector organizations are trying to overcome this issue on a priority basis. Suppose we want to locate or to track a vehicle in a metropolitan city. In this situation, surveillance cameras play an important role, but continuous recording videos of surveillance cameras generate a large amount of data and it is hard for an operator to analyze the data whenever any specific incident happens. Therefore, an intelligent video surveillance system's goal is to automatically monitor and analyze the data to help an operator in understanding the acquired video of a surveillance camera [1,2]. Those cameras observe changes in pixels due to the object crossing from the field of view of a camera. According to this scenario, it is only implemented in those areas where there is a rare movement of vehicles or in a parking area [1,3].

The basic principle of cameras in a surveillance system is to detect, associate, and track a vehicle/person across the range of cameras at a number of locations with a different background. Object (Re-Id) aims to address the surveillance problem by matching the query image with a large number of gallery images that are captured over a non-overlapping camera network. Locating and tracking a vehicle on overlapping and non-overlapping surveillance cameras can be performed on the vehicle Re-Id system. First, multi-camera tracking is the tracking an individual vehicle using multiple cameras; the identity of the vehicle has been retrieved from the second camera based on the information obtained from the first camera. Second, if the locations of the camera are known, based on the Re-Id system, it is possible to track the moving path of a vehicle from one point to another. Moreover, in surveillance and security applications, Re-Id models can be employed to track the suspicious vehicle in a crime scene. Moreover, real time surveillance system is shown in Figure 1.



**Figure 1.** A glance of the practical environment where vehicles are captured by cameras in the city.

Vehicle Re-Id gained more attention [4,5] in the last two years. Before that, computer vision field researchers were more focused on person Re-Id applications [6]. Person Re-Id and vehicle Re-Id are almost two similar problems, but vehicle Re-Id is more challenging as compared to person Re-Id, due to the lack of appropriate datasets and images of the same vehicle usually having an overlapped appearance even with large viewpoint variations. A human body is usually upright, and the texture or color of the dress will not vary severely in different viewpoints [7]. However, the visual features of vehicle in varying views is totally different, as shown in Figure 2. For instance, there is no overlapping on the visual pattern of the vehicle across the front side and rear viewpoints. Conventional person Re-Id methods, in which multi-view processing is not a large consideration, are not reasonable to be transferred directly to vehicle Re-Id [8].



**Figure 2.** Changing the viewpoint vehicle totally changes as compared to a person.

Due to demand towards vehicle Re-Id in public and private sectors, there are several approaches proposed but broadly divided into two categories. In the first one, vehicle ID, location, and time information are continuously transmitted towards the control center, and the global positioning system (GPS) is used for tracking the shipment. After that, it is used for so many fields like tracking vehicles, vehicle anti-theft management systems, and intelligent transportation systems. In the second category, the vehicle is identified when it comes in the field of view of the device terminal. Surveillance cameras, radio-frequency identification, and inductive loop lie in this category [9]. From all of these technologies, surveillance cameras are suitable for vehicle Re-Id [10], as the cameras do not need owner cooperation and also a camera that is already installed on the roadside is useful without any additional charges.

### 1.1. Issues Regarding Vehicle Re-Identification

Vehicle Re-Id is one of the interesting and demanding problems in the last two years in the computer vision research community. In order to track the vehicles from one frame of the camera to another, the frames of the different or same camera over a surveillance camera network is challenging. The reasons why vehicle tracking is challenging are discussed below.

- **Inter-class similarity:** This issue arises with similar model vehicles or different manufacturers of vehicles having a visually similar appearance, like two different makes, models, and types of vehicles having a similar rear or front side as shown in Figure 3 [11,12].



Figure 3. Inter-class similarity.

- **Intra-class variability:** In intra-class variability, same vehicle looks visually different over the same or different surveillance camera network due to an inconsistent environment [11,12].
- **Pose and viewpoint variations:** Vehicle pose and view angle varies due to the calibration of cameras on roadside. This is the reason for consistent changes in vehicle appearance and the vehicle looks totally different by changing view angle. In computer vision, sometimes two similar vehicles look more different in different viewpoints as compared to two different vehicles with the same view angle [8,13], as shown in Figure 4.

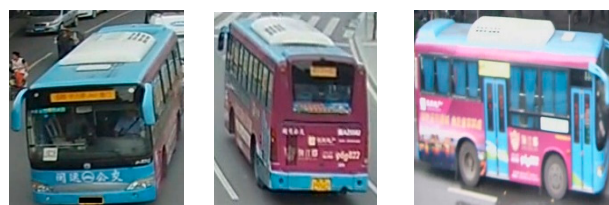


Figure 4. Effect of viewpoint variation on same vehicle.

- **Partial occlusions:** Some part of a targeted vehicle is covered by any other vehicle in a crowded scene or object, and as a result, the descriptor generated by an occluded image is corrupted [14,15].
- **Illumination changes:** Vehicle illumination changes camera to camera [12] or even on the same camera in a different time because of the inconsistent environment [16]. Vehicles' head and rear lights in the night have an effect on illumination; as a result, vehicle appearance also changes with a period of time and different camera networks [13,14,17].

- **Resolutions variation:** Resolution changes in the set of same vehicle images, due to camera calibration and other factors, as many old cameras are installed on roadsides that capture low-resolution images. Due to this factor, it is hard to differentiate between vehicle images with low discriminatory details.
- **Deformation:** Changes in the vehicle body, or shape due to an accident or load.
- **Background clutter:** This issue arises due to background color and vehicle's color being the same in an image.
- **Changes in color response:** Color is one of the important parameters in vehicle Re-Id but a camera to camera color response varies due to camera features and settings [16] and this may affect vehicle appearance.
- **Lighting effects:** Shadows and the specular reflection of a vehicle body adds noise to an image descriptor [11,18].
- **Long-term Re-Id:** With a large time and space between two views of the same vehicle there is a strong possibility that the vehicle may appear with some changes or be observed carrying some object/load through a different camera.

### 1.2. Applications

There are various practical applications of vehicle Re-Id systems; however, some major applications are discussed below.

- **Suspicious vehicle search:** Mostly, automobiles play a key role in criminal activities, such as when a terrorist uses it to attack different public places and soon leaves that place in the vehicle. However, it is a challenge for security officers to search suspicious vehicles within the limited time to stop criminals causing further loss.
- **Cross camera vehicle tracking:** In vehicle races, some viewers want to watch only one targeted vehicle, and when the vehicle comes in the field of view of the camera network at different times and places, then the broadcaster only focuses on that vehicle. With this vehicle Re-Id system, it is also helpful to track the specific vehicle path.
- **Automatic toll collection:** Vehicle Re-Id systems can also be applied on toll gates to automatically identify vehicle type and apply the toll rate.
- **Road access restriction management:** Mostly, in metropolitan cities of the world, heavy vehicles are not allowed to enter in the city at daytime, or some vehicles are not allowed on different days to run on roads or restricted areas where only officials or authorized vehicles can run.
- **Intelligent parking:** There are many sensitive places where this system can be applied at the entrance to identify the authorized vehicle and then allow it to park.
- **Traffic flow analysis:** To analyze the congestion on different routes at different times, such as peak hour calculations or behavior of the specific type of vehicle.
- **Vehicle counting:** This system is helpful in counting a specific type of vehicle.
- **Speed restriction management system:** Calculating the average velocity of a vehicle when it is passing by two consecutive camera locations.

The main contribution of our approach is to propose an appearance-based vehicle Re-Id model, in which we have applied a data augmentation technique on a dataset before training. Then we used the global channel to obtain a descriptor from the whole image, and also we used the local region channel to extract salient features from local cues with jointly incorporated attributes, such as color, type, model etc. Secondly, we propose a siamese neural network to verify the license plate in which two parallel convolutional neural networks (CNN) with the same structure and sharing the same weights in forward and backward computations are used.

## 2. Related Work

In order to provide context to the proposed method, this section is divided into two subheadings; both subheadings provide a concise and precise description of work related to vehicle Re-Id till now.

### 2.1. Appearance Attributes Based Vehicle Re-Identification

Vehicle Re-Id by visual appearance is a challenging task and regarded as fine-grained recognition [19], because the aim of the vehicle Re-Id task is to search specific vehicles over the surveillance camera network rather than a specific model, class, and type. Hence, extracted visual features should be more discriminative and capable of representing some salient identification marks, such as decoration stickers, annual inspection marks, etc. [12]. Methods exploiting vehicle Re-Id based on appearance are discussed in detail as follows.

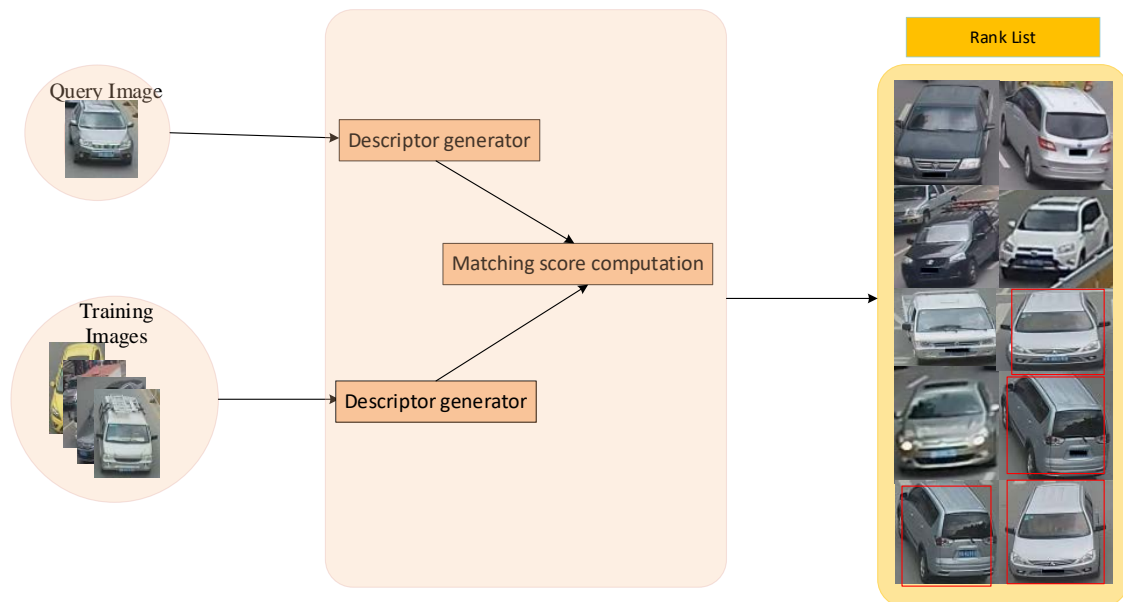
Due to the robust performance of the CNN in the recent past [20–22], it helps a lot in vehicle recognition and Re-Id tasks, such as verification [23], classification [24], and attribute prediction [25]. To analyze traffic using surveillance cameras and to obtain high accuracy is a most challenging task. Various vehicle Re-Id approaches are proposed to search the query image from a gallery set on bases of license plate recognition, color, and different attributes of the vehicle. Liu et al. [26] investigate some vehicle appearance features, like color, texture, and semantic attributes learned by the CNN. The authors also formulate a model by combining low-level features and high-level features to search targeted vehicles. Feris et al. [27] proposed system for vehicle detection and retrieval by attribute identification and vehicle color. Further, to reach the accurate targeted vehicle, the author uses recognized vehicle attributes. Wang et al. [28] proposed vehicle Re-Id architecture in which the system learns local and global features of vehicles and finally re-ranks results on the basis of a spatiotemporal regularization model.

Liu et al. [11] used a NULL spaced based fact model technique to filter the vehicle by appearance. The author's approach presents all three types of features, i.e., texture, color, and high-level attributes extracted from each training vehicle image, and combining them to obtain the unique (original) appearance feature to kernelize. Finally, research calculated the Euclidean distance of the input image with the gallery image in null space to measure the similarity for the final result. Jiang et al. [14] proposed an approach which consisted of a multi-branch architecture model. The multi-branch network consisted of two different blocks to learn labels: (1) A backbone network responsible for extracting appearance features consisting of a dense block. (2) A two-branch network responsible for extracting model and color features consisting of a coffee-net and inception block, respectively. Liu et al. [29] found that most researchers used global features and distance matrices for Re-Id, but with the appearance attributes based vehicle Re-Id technique, it is hard to re-identify the same vehicle because of intra-class variance and inter-class similarity. In comparison with the global appearance, local regions help: stickers attached to the front or rear glass, decorations, etc. To deal with Re-Id challenges, the authors proposed a region aware deep model (RAM). The RAM extracts the features from the local region with the addition of the global features. Shangzhi et al. [30] proposed scanning spatial and channel attention networks based on deep CNN. The attention network consists of two branches: (1) the spatial attention branch and (2) the channel attention branch. Respectively, the spatial attention branch is responsible for exploring more discriminatory regions of an input image. The channel attention branch is responsible for discovering the more discriminative channel in the network to improve network power. Both the channel attention network and spatial attention networks adjust weights and then combine the weights in an original feature map. Cui et al. [31] proposed a method in which they used a deep learning-based approach that is a two-branched multi-deep neural network fusion siamese neural network (MFSNN), and color, model, and annual inspection sticker or different marks pasted on windshields of vehicles are input to the network and calculate similarity between two vehicles. Zhang et al. [32] proposed the triplet-wise training of CNN architecture for vehicle Re-Id and they trained the network for feature extraction in the form of triplets (input, positive images, negative images), and after that, the CNN is divided into two streams to learn the feature vectors. A triplet stream gives constraints of triplet



loss and a classification stream gives the oriented loss. Despite all the above-discussed approaches, vehicle Re-Id was not well explored and is still suffering.

Figure 5 shows the standard vehicle Re-Id framework. After image feature vector generation for the input probe and each gallery image, the framework calculates the matching score between the gallery and probe images and in the last ranked list generated on the basis of the computed score.



**Figure 5.** Standard vehicle re-identification system.

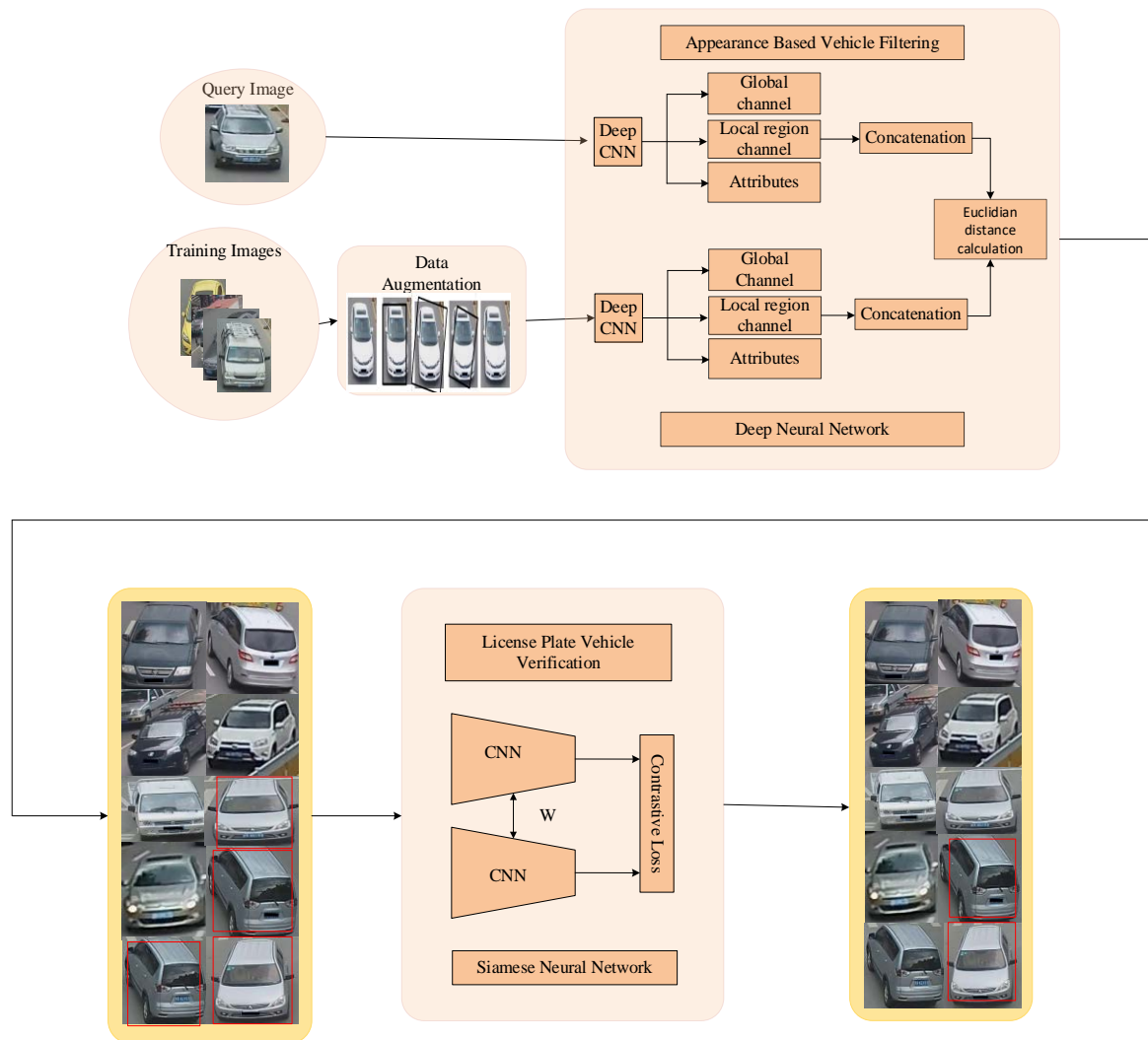
## 2.2. License Plate Based Vehicle Re-Identification

The license plate number is the conventional approach to identify that either a vehicle is same or not [33]. License plates provide unique information of vehicles [12]. An automatic license plate recognition system is useful to re-identify a vehicle by matching characters of the license plate image. License plate recognition systems process first extracts of the license plate region from the vehicle image to convert characters into a string and then compares the two strings to re-identify the vehicle. Since license plate recognition systems cannot give 100% accuracy and misread some of the characters due to unconstrained surveillance environments, other issues like illumination, various viewpoints, and also image resolution affect the accuracy of license plate recognition systems. Various previously proposed approaches are discussed below.

Qadri et al. [33] proposed a system for vehicle Re-Id using number plates in which a vehicle was detected in the image first and then the vehicle number plate was extracted through a smearing algorithm. After that, the next step was to recognize the characters written on the number plate; the authors used an optical character recognition technique and lastly compared characters with the stored database to get particular vehicle information, such as registration address, vehicle owner, etc. Pham et al. [34] proposed a CNN-based method for vehicle license plate recognition, firstly preprocessing the input image of vehicle. Secondly, the authors applied different techniques like filtering, and thresholding for segmentation, and finally, for character recognition, they trained the CNN classifier. Watcharapinchai et al. [35] have evaluated several comparative techniques for vehicle license plate recognition and the authors found that the generalized edit distance (GED) achieved the highest precision. On the other hand, 2-Grams had the highest recall, but for best performance, GED and 2-Grams should have the highest value, so they proposed the approach by combining both techniques and achieved a satisfactory result.

### 3. Overview of the Proposed Method

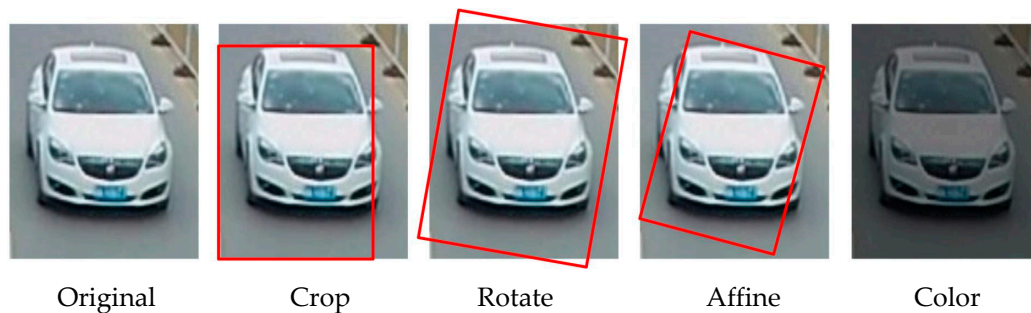
The architecture and working procedure of the proposed deep model is shown in Figure 6. In the proposed deep model, we first input the query image of a vehicle from a surveillance camera to the Re-Id system; the system then searches the same vehicle with two steps. In the first phase, the system searches similar vehicles from the gallery set on the basis of vehicle appearance. These extracted small number of vehicles in the second phase are fed for license plate verification. For this purpose, the siamese neural network calculates the distance between license plates of the query image and the gallery images. With this bird-eye-view of our methodology, the following gives greater detail.



**Figure 6.** Illustration of the complete architecture of our proposed method. Firstly, we input the probe image of the vehicle to the Re-Id system from surveillance cameras. Our proposed system obtains the specific vehicle in two steps: (1) search the query vehicle image from the gallery set on the bases of global and local appearance with vehicle attributes; (2) license plate verification between the query image and the obtained gallery images after step one to get the targeted vehicle.

In our approach, the vehicle Re-Id system dataset is divided into two parts: training and testing. The Re-Id system learns and sets parameters on training data, and performance is measured on testing data. It is worth mentioning that during the training process, problems occur when the training dataset consists of correlated data which cause over fitting due to system parameters getting over-tuned on specific viewpoints, illumination, etc. Therefore, the Re-Id system does not form well on other

viewpoints, illumination, etc., on which it is not trained. For example, usually available dataset images are captured by a small number of cameras. As a consequence, most of the images have the same angle, color, and background, and therefore influence increases and the system starts recognizing these feature descriptors. To address this problem, we first applied data augmentation on the training set to increase the variability of the available training dataset and to compensate for pose, lighting conditions, and partial occlusions, as shown in Figure 7. This enhances the performance of vehicle Re-Id system.



**Figure 7.** Data augmentation based on linear transformation.

The data augmentation performed in this phase is part of appearance-based Re-Id. Further steps of this process are mentioned in the following subsection.

### 3.1. Appearance Attributes Based Vehicle Re-Identification

Existing methods of vehicle Re-Id usually extract the global appearance features. However, extracting only global features for vehicle Re-Id is hard to differentiate in the case of different vehicles of same make and model because of a similar global appearance look. Therefore, global features as well as local features are important, such as descriptors generated from annual inspection stickers, logos and different marks and decoration paints on the windshield glass which help to mitigate the negative effects of the intra-class variance and inter-class similarity. Whereas local cues can improve the discrimination and robustness of feature representation. Hence, in the proposed method, we use both global and local cues to generate better discriminative representations of vehicles.

The proposed model consists of a shared CNN with three channels, i.e., global channel, local region channel, and attribute channel. Every channel's generated discriminative features are trained with a separate classifier and then the feature vectors are concatenated for original vehicle Re-Id. To get a small number of vehicles from the whole dataset, training dataset feature vectors are matched with the query image feature vector. Using the global channel, we can extract a single feature vector that generalizes the whole image, and using the local region channel, we can extract features from multiple points, like annual inspection stickers, logos, decorations, and different marks on the windshield of the vehicle image, to make more robust attributes for learning the vehicles' intrinsic features, containing vehicle model, color, and type information. For example, different vehicles with a similar global appearance can be seen in Figure 8, and each column shows two different vehicles. The differences on local regions are highlighted with red circles. It can be observed that the differences between similar vehicles mostly lie in some local regions.





**Figure 8.** Vehicles that are the same in global appearance but differentiated by local regions, as marked with red circles.

Based on small number of vehicles achieved in the first phase, we perform license plate based vehicle Re-Id in the following subsection.

### 3.2. License Plate Based Vehicle Re-Identification

License plate recognition (LPR) systems read license plate characters to match between the query vehicle and the gallery images to Re-Id the vehicle. However, for this purpose, it is important to have a consistent environment because it requires auxiliary infrastructure, such as flashlights and sensors. In fact, license plates are often available in inconsistent environments, which makes it difficult because of lighting, viewpoint angle, and occlusion. Hence, we are here verifying vehicle license plates instead of recognizing if two license plates are the same or not, using the siamese neural network (SNN). SNN was proposed by Chopra et al. [36] and was designed for person face verification. SNN is suitable for problems where a number of categories are large but instances from each category are limited. SNN is designed with two identical parallel convolutional neural networks to extract the robust discriminative feature vectors from license plate pairs. Convolutional neural networks are trained in an end-to-end way to map pixel images to outputs. The key benefit of using the CNN in our approach is that we can learn optimal shift-invariant local feature detectors and generate features that are robust to geometric distortions of the query image and then calculate the Euclidean distance between the feature vectors. In a license plate verification system, distance should be the minimum between a pair of license plates from the same vehicle, and large from pairs of license plates from different vehicles.

Suppose  $A_1$  and  $A_2$  are input pairs of license plates,  $B$  is a binary label of the pair. The pair  $= 0$  if the images  $A_1$  and  $A_2$  belong to the same vehicle and  $B = 1$  otherwise,  $W$  the weights of the convolutional neural network, and  $C_w(A_1)$  and  $C_w(A_2)$  are features extracted from  $A_1$  and  $A_2$  images of the vehicle license plates, respectively. Then our system calculates the Euclidean distance between them as:

$$E_w(A_1, A_2) = ||C_w(A_1) - C_w(A_2)||.$$

With  $E_w(A_1, A_2)$ , the contrastive loss is defined as:

$$L(W, (A_1, A_2, B)^i) = (1 - B) \cdot L_c(E_w(A_1, A_2)^i) + B L_I(E_w(A_1, A_2)^i)$$

where  $(B, A_1, A_2)^i$  is the  $i$ -th tuple of two training vehicle license plates.  $L_c$  is the partial loss function for the same license plate and  $L_i$  is the loss function for different license plates.

#### 4. Experiment and Analysis

To measure the efficiency of the vehicle Re-Id system, comprehensive experiments were conducted on our proposed deep model and compared with state-of-the-art approaches. Moreover, to evaluate the approach, it was necessary that the dataset should represent real-world camera surveillance data that contains unavoidable factors like occlusion, changes in illumination, and background clutter, etc. That is why here we are using the VeRi [26] benchmark dataset of vehicle Re-Id. VeRi is a large-scale benchmark dataset mostly used for vehicle Re-Id. It is prepared in a real-world scenario with surveillance cameras. The dataset consists of 776 vehicles with 50,000 total images captured from 20 different cameras, and each vehicle is captured by 2 to 18 cameras with different illumination, resolution, and occlusion. To re-identify complex vehicle models, dataset images are labeled with different attributes, such as vehicle brand name, color, type, etc.

##### 4.1. Implementation Details

We have trained the framework on a NVIDIA GeForce GTX 1080 GPU, 8 GB RAM memory, Ubuntu 16.04 LTS operating system with Core i7 Intel CPU. To accelerate the experiment, Graphics processing unit (GPU) have been utilized. We developed the whole framework in Python and most of the time the libraries used were NumPy, matplotlib, scikit-learn, and SciPy.

First, we applied an appearance-based approach on the VeRi dataset and then examined the license plate verification for specific vehicle search. A shared multiple layer convolutional neural network was used to generate feature map  $N$  and then feature map  $N$  was passed to the global channel, the region channel, and the attributes channel to generate more discriminative features. The global channel pooled the feature maps  $N$  and the fully connected (FC) layers generated global features that represented the whole image. The region channel split feature maps  $N$  into  $m$  local regions of the vehicle, after each local region part we embedded the pooling layer and the fully connected layers to generate feature vectors of each local part. Lastly, the fully connected layer of the global channel was passed to the attribute channel and attribute features were generated by fully connected layers; ResNet 152 pre-trained on ImageNet was used to classify features that were extracted from the global channel, the local channel, and the attributes channel. We have used the siamese neural network (SNN) as a feature extractor for license plates and before training all the license plate number pairs were shuffled. Our SNN consisted of two identical convolutional networks and a contrastive loss function. The input to the system was a pair of vehicle license plate images. The query license plate images were passed through the sub-CNN networks, yielding two outputs which were passed to the loss function. The implementation of the SNN structure was in the TensorFlow deep learning tool.

##### 4.2. Results

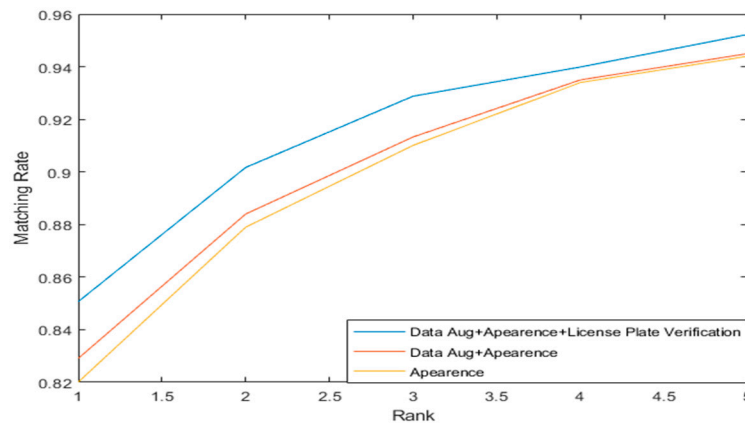
Figure 9 presents results of our proposed method based on appearance, data augmentation with appearance, and data augmentation with appearance with license plate verification. Data augmentation and data augmentation with license plate verification improved the mean average precision ( $mAP$ ) by 0.75% and 3.37% compared to appearance results, respectively. We have used  $mAP$  and cumulative matching characteristics (CMC) for evaluation of the technique; the  $mAP$  metric evaluates the overall performance for Re-Id.  $mAP$  was calculated for each probe image as follows:

$$AP = \sum_{k=1}^n p(k) * f(K) / n_{total}$$

where  $n$  denotes the total number of entries in the result list,  $p(k)$  is precision,  $f(k)$  is a function, it is equal to 1 if the result is correct, otherwise 0.  $N_{total}$  is number of vehicles images in test set that are same as a probe vehicle image.

$$mAP = \sum_{t=1}^t \frac{AP(t)}{T}$$

where  $T$  denotes the size of input set.



**Figure 9.** Cumulative matching characteristics (CMC) on the VeRi-776 dataset.

CMC calculates the probability of the input vehicle image appearing in different lengths of the result list. Meanwhile, the effectiveness of the contribution can also be seen in CMC curves, as shown in Figure 9.

As shown in Figure 9, our proposed complete architecture with license plate verification achieved the best performance.

#### 4.3. Performance Comparison with State-of-the-Art Methods

Performance comparison of multiple state-of-the-art approaches including PROVID [11], MICRO [16], Light vgg m+SCAN [30], VRSDNet [37], RAM [29], SCCN-Ft+CLBL-8-Ft [8], and GS-TRE loss W/mean VGGM [12] with our proposed approach on the VeRi dataset is shown in Table 1. The VeRi dataset was chosen for comparison because it consists of supplementary viewpoints, different illumination, and resolutions as compared to other datasets. The VeRi dataset fulfills all the aspects of real-world camera surveillance data for vehicle-related applications.

**Table 1.** Comparison with some state-of-the-art approaches. *mAP*: mean average precision.

Methods	<i>mAP</i>	HIT-1%	HIT-5%
PROVID [11]	53.42	81.56	95.11
FACT [26]	18.75	52.21	72.88
GoogLeNet [23]	17.89	52.32	72.17
VGG [38]	12.76	44.10	62.63
FACT + Plate-SNN + STR [39]	27.77	61.44	78.78
GS-TRE loss W/mean VGGM [12]	59.47	96.24	98.97
JFSDL [40]	53.53	82.90	91.60
ABLN-Ft-16 [41]	24.92	60.49	77.33
SCCN-Ft+CLBL-8-Ft [8]	25.12	60.83	78.55
Appearance + Color + Model + Re-Ranking [14]	61.11	89.27	94.76
Siamese-CNN-Path-LSTM [42]	58.27	83.49	90.04
RAM [29]	61.5	88.6	94.0
OIF+ST [28]	51.4	92.35	...
MTCRO [16]	62.61	87.96	94.63
ResNet-50 + GoogleNet + Feature fusion via correlative sparse ranking [43]	58.21	90.52	93.38
Light_vgg_m+SCAN [30]	49.87	82.24	90.76
VAMI [13]	50.13	77.03	90.82
XVGAN [44]	24.65	60.20	77.03
VGG+C+T [19]	58.78	86.41	92.91
VRSDNet [37]	53.45	83.49	92.55
QD-DLF [45]	61.83	88.50	94.46
Combining Network [46]	33.78	60.19	77.40
Appearance (Ours)	59.43	82.01	94.40
Data Augmentation + Appearance (Ours)	60.21	82.91	94.50
Data Augmentation + Appearance + License plate verification (Ours)	62.80	85.07	95.23

Firstly, from Table 1, it is observed that our method outperforms on the VeRi dataset as compared to all recent methods. It achieves 62.80% in mAP, 85.07 HIT@1 and 95.23 HIT@5. The VeRi dataset contains spatio-temporal information and most of the methods [11,13,16,28,39] exploit spatio-temporal information in the vehicle Re-Id system. It is worth mentioning here that our approach outperforms without using any spatio-temporal information, whereas Orientation Invariant Feature Embedding (OIFE) [28] needs additional annotations of key points during training for their orientation invariant embedded learning. It is noticeable in our approach that the training procedure was more convenient than many other methods presented in Table 1.

## 5. Conclusions

In this paper, we presented a deep convolutional network learning-based approach for the vehicle re-identification (Re-Id) problem. Re-Id is one of the important and challenging areas in intelligent transportation systems. Despite high significance, it is not well explored as compared to the similar problem that is person Re-Id. Our deep model extracts the global and local visual features that are a more robust representation of the vehicle image. The attributes channel extracts the color, model, and type of vehicle. In addition, we utilized a license plate verification technique using the siamese neural network between shortlisted vehicles by appearance and the targeted vehicle. The results on the VeRi-776 dataset show the robustness of our method.

In future, we aim to explore possibilities to enhance the overall vehicle Re-Id. There is significant potential to extend the approach with an attention network for salient feature extraction.

**Author Contributions:** All the authors made significant contributions to this work. Introduction, Z. and M.U.A.; Related work, Z. and M.U.A.; Methodology, Z., M.S.K., and J.D.; Results, Z., R.K., and M.S.K.; Supervision, J.C.; Review and Editing, J.D.

**Funding:** This paper has been supported by the project of NSFC (ID: 61602098), and by the Funding of zhongyanggaoxiao ZYGX2018J075 and also by the project of 18CYRC0017 of Sichuan science and technology support plan.

**Acknowledgments:** We are thankful to anonymous editor and reviewers for their valuable comments and kind suggestions in early versions of this article.

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

## References

1. Makris, D.; Ellis, T.; Black, J. Bridging the gaps between cameras. *CVPR* **2004**, *2*, 205–210.
2. Tieu, K.; Dalley, G.; Grimson, W.E.L. Inference of non-overlapping camera network topology by measuring statistical dependence. *Proc. IEEE Int. Conf. Comput. Vis.* **2005**, *II*, 1842–1849.
3. Zelniker, E.E.; Gong, S.; Xiang, T. Global abnormal behaviour detection using a network of CCTV cameras. In Proceedings of the The Eighth International Workshop on Visual Surveillance-VS2008, Marseille, France, 17 October 2008.
4. Huang, C.L.; Liao, W.C. A vision-based vehicle identification system. *Proc. Int. Conf. Pattern Recognit.* **2004**, *4*, 364–367.
5. Hoffman, A.J.; Pretorius, A.; Wang, Y. Geometry Based Analysis of an Alternative RFID Interrogator Placement for Electronic Vehicle Identification. In Proceedings of the IEEE Conference on Intelligent Transportation Systems, Las Palmas, Spain, 15–18 September 2015; pp. 2390–2397.
6. Zheng, L.; Yang, Y.; Hauptmann, A.G. Person Re-identification: Past, Present and Future. *arXiv*, 2016; arXiv:1610.02984v1.
7. Farenzena, M.; Bazzani, L.; Perina, A.; Murino, V.; Cristani, M. Person re-identification by symmetry-driven accumulation of local features. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 2360–2367.
8. Zhou, Y.; Liu, L.; Shao, L. Vehicle Re-Identification by Deep Hidden Multi-View Inference. *IEEE Trans. Image Process.* **2018**, *27*, 3275–3287. [[CrossRef](#)] [[PubMed](#)]
9. Jeng, S.-T. *Real Time Re-Identification System for Freeway Performance Measurement*; UC Berkeley: Berkeley, CA, USA, 2007.

10. Ji, X.; Zhang, G.; Chen, X.; Guo, Q. Multi-Perspective Tracking for Intelligent Vehicle. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 518–529. [[CrossRef](#)]
11. Liu, X.; Liu, W.; Mei, T.; Ma, H. PROVID: Progressive and Multimodal Vehicle Reidentification for Large-Scale Urban Surveillance. *IEEE Trans. Multimed.* **2018**, *20*, 645–658. [[CrossRef](#)]
12. Bai, Y.; Lou, Y.; Gao, F.; Wang, S.; Wu, Y.; Duan, L.Y. Group-sensitive triplet embedding for vehicle reidentification. *IEEE Trans. Multimed.* **2018**, *20*, 2385–2399. [[CrossRef](#)]
13. Zhou, Y.; Shao, L.; Dhabhi, A. Viewpoint-aware Attentive Multi-view Inference for Vehicle Re-identification. *CVPR* **2018**, 6489–6498.
14. Jiang, N.; Xu, Y.; Zhou, Z.; Wu, W. Multi-Attribute Driven Vehicle Re-Identification with Spatial-Temporal Re-Ranking. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 858–862.
15. Zhu, J.; Zeng, H.; Lei, Z.; Liao, S.; Zheng, L.; Cai, C. A Shortly and Densely Connected Convolutional Neural Network for Vehicle Re-identification. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 3285–3290.
16. Xu, D.; Wang, T. A Framework With A Multi-task CNN Model Joint With A Re-ranking Method For Vehicle Re-identification. In Proceedings of the the 10th International Conference on Internet Multimedia Computing and Service, Nanjing, China, 17–19 August 2018; Volume 1.
17. Frías-Velázquez, A.; Van Hese, P.; Pižurica, A.; Philips, W. Split-and-match: A Bayesian framework for vehicle re-identification in road tunnels. *Eng. Appl. Artif. Intell.* **2015**, *45*, 220–233. [[CrossRef](#)]
18. Gazzah, S.; Essoukri, N.; Amara, B. Vehicle Re-identification in Camera Networks: A Review and New Perspectives. In Proceedings of the ACIT'2017 The International Arab Conference on Information Technology, Yasmine Hammamet, Tunisia, 22–24 December 2017.
19. Zapletal, D.; Herout, A. Vehicle Re-Identification for Automatic Video Traffic Surveillance. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Las Vegas, NV, USA, 26 June–1 July 2016.
20. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the NIPS 2012: Neural Information Processing Systems Conference, Lake Tahoe, NV, USA, 3–6 December 2012.
21. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv*, 2014; arXiv:1409.1556.
22. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
23. Yang, L.; Luo, P.; Loy, C.C.; Tang, X. A Large-Scale Car Dataset for Fine-Grained Categorization and Verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; Volume 1.
24. Krause, J.; Stark, M.; Deng, J.; Fei-Fei, L. 3D object representations for fine-grained categorization. In Proceedings of the the IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 2–8 December 2013; pp. 554–561.
25. Xiang, Y.; Choi, W.; Lin, Y.; Savarese, S. Data-Driven 3D Voxel Patterns for Object Category Recognition. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.
26. Liu, X.; Liu, W.; Ma, H.; Fu, H. Large-Scale Vehicle Re-Identification in urban Surveillance Videos. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Seattle, WA, USA, 11–15 July 2016.
27. Feris, R.S.; Siddiquie, B.; Petterson, J.; Zhai, Y.; Datta, A.; Brown, L.M.; Pankanti, S. Large-scale vehicle detection, indexing, and search in urban surveillance videos. *IEEE Trans. Multimed.* **2012**, *14*, 28–42. [[CrossRef](#)]
28. Wang, Z.; Tang, L.; Liu, X.; Yao, Z.; Yi, S.; Shao, J.; Yan, J.; Wang, S.; Li, H.; Wang, X. Orientation Invariant Feature Embedding and Spatial Temporal Regularization for Vehicle Re-identification. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 379–387.



29. Liu, X.; Zhang, S.; Huang, Q.; Gao, W. RAM: A Region-Aware Deep Model for Vehicle Re-Identification. In *Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME)*, San Diego, CA, USA, 23–27 July 2018; pp. 1–6.
30. Teng, S.; Liu, X.; Zhang, S.; Huang, Q. SCAN: Spatial and Channel Attention Network for Vehicle Re-Identification. In *Proceedings of the Advances in Multimedia Information PCM*; Springer International Publishing: Cham, Switzerland, 2018; pp. 350–361.
31. Cui, C.; Sang, N.; Gao, C.; Zou, L. Vehicle re-identification by fusing multiple deep neural networks. In *Proceedings of the 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Montreal, QC, Canada, 28 November–1 December 2017; pp. 1–6.
32. Zhang, Y.; Liu, D.; Zha, Z.J. Improving triplet-wise training of convolutional neural network for vehicle re-identification. In *Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME)*, Hong Kong, China, 10–14 July 2017; pp. 1386–1391.
33. Qadri, M.T.; Asif, M. Automatic Number Plate Recognition System for Vehicle Identification using Optical Character Recognition. In *Proceedings of the International Conference on Education Technology and Computer AUTOMATIC*, Singapore, 17–20 April 2009; pp. 3–6.
34. Pham, V.H. CNN-Based Character Recognition for License Plate Recognition System. In *Asian Conference on Intelligent Information and Database Systems*; Springer: Cham, Switzerland, 2018; pp. 594–603.
35. Watcharapinchai, N.; Rujikietgumjorn, S. Approximate license plate string matching for vehicle re-identification. In *Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Lecce, Italy, 29 August–1 September 2017; pp. 2–7.
36. Chopra, S.; Hadsell, R.; LeCun, Y. Learning a Similarity Metric Discriminatively, with Application to Face Verification. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA, 20–25 June 2005.
37. Zhu, J.; Du, Y.; Hu, Y.; Zheng, L.; Cai, C. VRSDNet: Vehicle re-identification with a shortly and densely connected convolutional neural network. In *Multimed. Tools Applications*; Springer: Hoboken, NJ, USA, 2018.
38. Liu, H.; Tian, Y.; Wang, Y.; Pang, L.; Huang, T. Deep Relative Distance Learning: Tell the Difference Between Similar Vehicles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27–30 June 2016.
39. Liu, X.; Liu, W.; Mei, T.; Ma, H. A Deep Learning-Based Approach to Progressive Vehicle Re-identification for Urban Surveillance. In *Proceedings of the European Conference on Computer Vision*, Amsterdam, The Netherlands, 11–14 October 2016; Volume 1, pp. 869–884.
40. Zhu, J.; Zeng, H.; Du, Y.; Lei, Z.; Zheng, L.; Cai, C. Joint feature and similarity deep learning for vehicle re-identification. *IEEE Access* **2018**, *6*, 43724–43731. [[CrossRef](#)]
41. Zhou, Y.; Shao, L. Vehicle re-identification by adversarial bi-directional LSTM network. In *Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Tahoe, NV, USA, 12–15 March 2018; pp. 653–662.
42. Shen, Y.; Xiao, T.; Li, H.; Yi, S.; Wang, X. Learning Deep Neural Networks for Vehicle Re-ID with Visual-spatio-temporal Path Proposals. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 22–29 October 2017.
43. Sun, D.; Liu, L.; Zheng, A.; Jiang, B.; Luo, B. Visual Cognition Inspired Vehicle Re-identification via Correlative Sparse Ranking with Multi-view Deep Features. In *Advances in Brain Inspired Cognitive Systems*; Springer International Publishing: Cham, Switzerland, 2018.
44. Zhou, Y.; Shao, L. Cross-View GAN Based Vehicle Generation for Re-identification. *BMVC* **2017**, *1*, 1–12.
45. Zhu, J.; Zeng, H.; Huang, J.; Liao, S.; Lei, Z.; Cai, C.; Zheng, L. Vehicle Re-identification Using Quadruple Directional Deep Learning Features. *arXiv*, 2018; arXiv:1811.05163v1.
46. Tang, Y.; Wu, D.; Jin, Z.; Zou, W.; Li, X. Multi-Model Metric Learning for Vehicle Re-Identification in Traffic Surveillance Environment. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Beijing, China, 17–20 September 2017.

