Review

Trends in Vehicle Re-Identification Past, Present, and Future: A Comprehensive Review

Zakria [1], Jianhua Deng [1,*], Yang Hao [2,*], Muhammad Saddam Khokhar [3], Rajesh Kumar [4], Jingye Cai [1], Jay Kumar [4] and Muhammad Umar Aftab [5]

1. School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China; zakria@uestc.edu.cn or zakria.uestc@hotmail.com (Z.); jycai@uestc.edu.cn (J.C.)
2. Institute of Applied Electronic (IAE), China Academy of Engineering Physics, Mianyang 621900, China
3. School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212003, China; muhammadsaddam@stmail.ujs.edu.cn or saddam_khokhar@hotmail.com
4. Yangtze Delta Region Institute (Huzhou), University of Electronic Science and Technology of China, Huzhou 313001, China; raja@uestc.edu.cn (R.K.); jay@std.uestc.edu.cn (J.K.)
5. Department of Computer Science, National University of Computer and Emerging Sciences, Chiniot-Faisalabad Campus, Chiniot 35400, Pakistan; Umar.aftab@nu.edu.pk
* Correspondence: jianhua.deng@uestc.edu.cn (J.D.); yhao@caep.cn (Y.H.)

Abstract: Vehicle Re-identification (re-id) over surveillance camera network with non-overlapping field of view is an exciting and challenging task in intelligent transportation systems (ITS). Due to its versatile applicability in metropolitan cities, it gained significant attention. Vehicle re-id matches targeted vehicle over non-overlapping views in multiple camera network. However, it becomes more difficult due to inter-class similarity, intra-class variability, viewpoint changes, and spatio-temporal uncertainty. In order to draw a detailed picture of vehicle re-id research, this paper gives a comprehensive description of the various vehicle re-id technologies, applicability, datasets, and a brief comparison of different methodologies. Our paper specifically focuses on vision-based vehicle re-id approaches, including vehicle appearance, license plate, and spatio-temporal characteristics. In addition, we explore the main challenges as well as a variety of applications in different domains. Lastly, a detailed comparison of current state-of-the-art methods performances over VeRi-776 and VehicleID datasets is summarized with future directions. We aim to facilitate future research by reviewing the work being done on vehicle re-id till to date.

1. Introduction

Due to growing global population, commercial activities have been extensively increasing, which leads everyone to access road transportation as a source of mobility. Due to easy accessibility of road transportation system, traffic on roads is massively increasing that not only creates the problem of high traffic congestion but also a drastic increase in carbon dioxide emissions. Along with these issues, road accident risks and the overall transportation complexity increases as well. Therefore, a smooth transportation source and medium is always required for growing commercial activities. Furthermore, traffic management authorities are facing hectic challenges to maintain an undisturbed transportation system. Their task includes tracking the suspicious vehicle, handling traffic jam, and to check whether the vehicle is registered or not. Maintaining undisturbed transportation becomes harder when a large number of vehicles are on the roads.

1.1. Intelligent Transportation System

Transport is essential for the daily routine functioning of the economy and the society. Over the past few decades there is huge development, deployment, and growth in the

transport system and have notable effect of development in society and daily life. Therefore, transportation should be redefined as ITS. Currently, not only mechanical and engineering fields are doing research and development for better transportation facility, but computer science related concepts are also playing major role for instance, artificial intelligence (AI), communication, machine learning (ML), internet and so many other emerging technologies.

Due to traffic problems in China, the average speed of vehicle has been decreased to 20 km/h, even in some areas between 7 and 8 km/h [1,2]. Such low speed of vehicles for a long time on roads is a threat for the natural environment of the world like exhaust emissions that deteriorate air quality. In order to deal traffic problems and alleviate the pressure of vehicles on roads, the governments are investing too much on research and ITS development. ITS based infrastructure strengthens the relationship between people, vehicles, and road networks.

ITS have the capability to enhance the performance of current transportation system and make it efficient, safe, comfortable as well as reduces harmful environmental consequences. ITS based real-time applications include electronic payment systems, traffic management systems, emergency vehicle pre-emption management system, advanced vehicle control systems, weather precautionary measures management system, and commercial vehicle operations. Applications of ITS now regularly deployed, such as closed-circuit television surveillance, automatic car parking, electronic toll collection, border control, and in-car navigation equipment. Therefore, an ITS is needed to analyze the recorded video, control, maintain and communicate to ground transport and improve mobility and manage problems efficiently. Furthermore, Figure 1 demonstrates the ITS based environment.
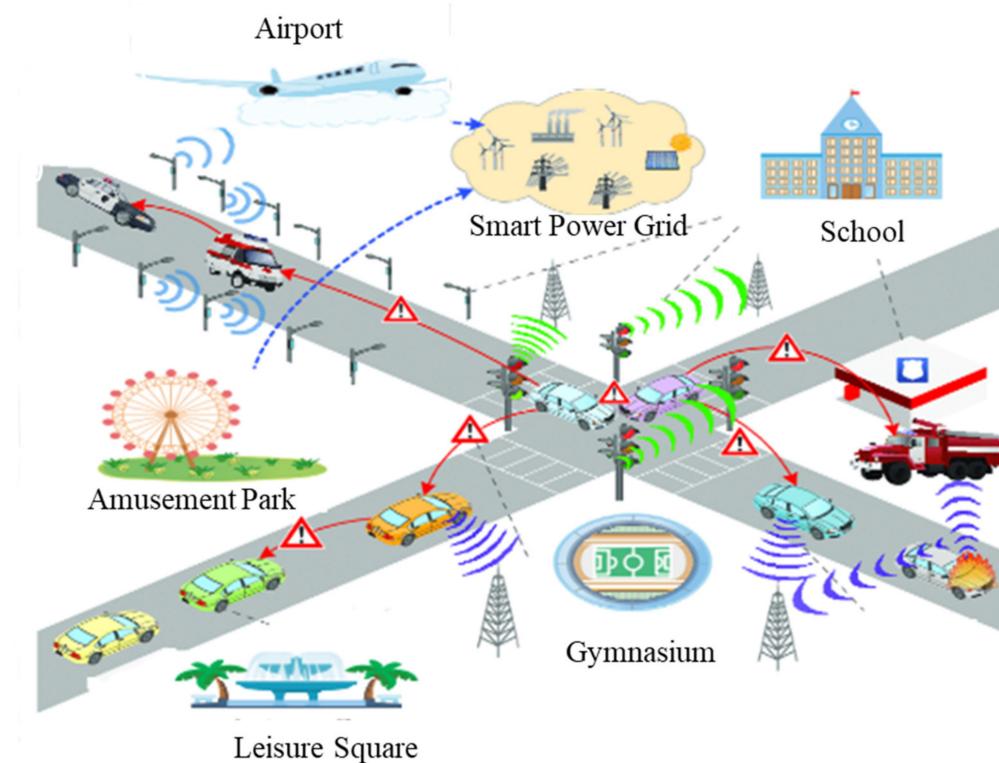
Figure 1. Depicts smart city and intelligent transportation system.

1.2. Video Surveillance

In metropolitan cities, cameras are widely adopted in numerous areas to monitor activities [3]; but most of the current video surveillance systems provides the facilities like capture, storage and distribute video, while leaves unwanted event detection task totally on human operators. Human operator-based monitoring of the surveillance system is not as efficient and a very labour-intensive task, as shown in Figure 2. It requires full visual attention by watching the video in control room and it is very difficult for single person

as everyday tasks. Specifically, the ability to focus and react to occasionally occurring activities that require full attention. Furthermore, millions of hours of video data generated by multiple cameras over surveillance network require large number of operators for the task. It's almost infeasible, inefficient and costly to obtain real-time prevention.

Figure 2. Shows view of manually traffic monitoring at control room.

Due to digital cameras and the advent of powerful computing resources, automatic video analysis become possible and more and more common in video surveillance applications [4], thus reduces the labor cost. Practically, the objective of automatic video analysis for safety, security, and surveillance is to detect automatically unwanted events or situations that need security attention. Automated video analysis not only process the data faster but also significantly improve the ability to preempt incidents on time. Augmenting security staff with automatic processing increases their efficiency and effectiveness. For the posterior mode, searching a specific vehicle in hundreds of hours of camera recorded video footage needs large number of officers to do this task and takes a lot of time. Automated content-based video retrieval reproducing and assisting human analysis on recorded videos largely enhances forensic capabilities. Furthermore, the surveillance systems application's main goal is to develop intelligent systems that automate the human decision-making mechanism.

An important task to maintain a smooth transport system is to re-identify the specific vehicle that appeared in different cameras over the surveillance network. The vehicle re-id module in ITS should recognize same vehicle that appears in surveillance cameras installed in different geographical locations. Specifically, vehicle re-id can be treated as a fine-grained recognition problem [5,6] that identifies the subordinate type of input class. However, the vehicle re-id problem's granularity is much finer since the system should search specific targeted vehicle instead of the same vehicle model and type. Moreover, recently vehicle re-id gained more attention in research community because of various significant real-world applications. It is a difficult task to analyze the surveillance environment for effective vehicle identification. An example of practical environment can be seen in Figure 3, where surveillance cameras can be observed over roads and public places.
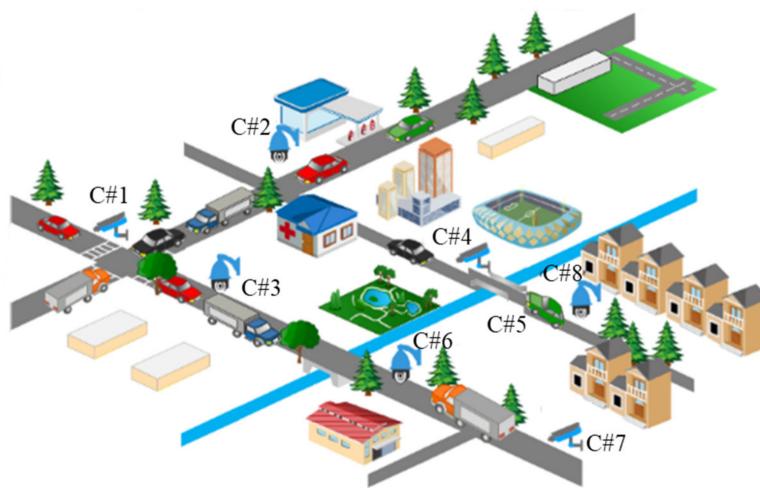
Figure 3. Illustrates the practical scenario of surveillance camera network.

1.3. Re-Identification

In a surveillance camera without overlapping vision, re-id is defined as a task to identify objects' captured images taken from different camera networks. It is used to know whether the object image captured by multiple surveillance cameras matches the same object or a different image of the object. Object re-id technology has a significant role in multi-object tracking, intelligent monitoring, and other fields. Recently, re-id gained extensive attention in the computer vision research community. The main application fields of an object re-id are vehicle re-id and person re-id.

Formally, re-id can be defined as a matching task. A targeted image (Query) is matched against a gallery set image (representing the previously captured images in the surveillance camera network). Thus, the query of re-identifying targeted image can be defined by its descriptor P, and it is formulated as:

$$T = arg_{T_i} \min D\ (T_i, Q), T_i \in \mathcal{T} \tag{1}$$

where $\mathcal{T} = \{T_1, \dots, T_N\}$ is a gallery set of N image descriptors, and $D(,)$ represents the distance metric. Therefore, to solve above the re-id problem, it is important first to answer how we can represent targeted object using a descriptor to robust performance. Furthermore, rest of the paper investigates this topic.

Vehicle Re-identification: Similar to person re-id, vehicle re-id is also a demanding task in camera surveillance. Aim of vehicle re-id is to match vehicle images with already captured vehicle images over the camera network [7–9]. However, due to surveillance cameras on the roads for smart cities and traffic management, the demand to perform vehicle search from the gallery set is increased. Vehicle re-id is similar to several other applications, such as person re-id [10], behavior analysis [11], cross-camera tracking [12], vehicle classification [13], object retrieval [14], object recognition [15,16], and so on.

To understand designing the vehicle re-id system, we analyze how a person re-identifies the vehicle. A person re-identifies vehicle by keeping in mind some characteristics like unique feature, color, size etc., our brain and eyes are learned to detect and identify different objects, as shown in Figure 4 and how system identify vehicle is shown in Figure 5.
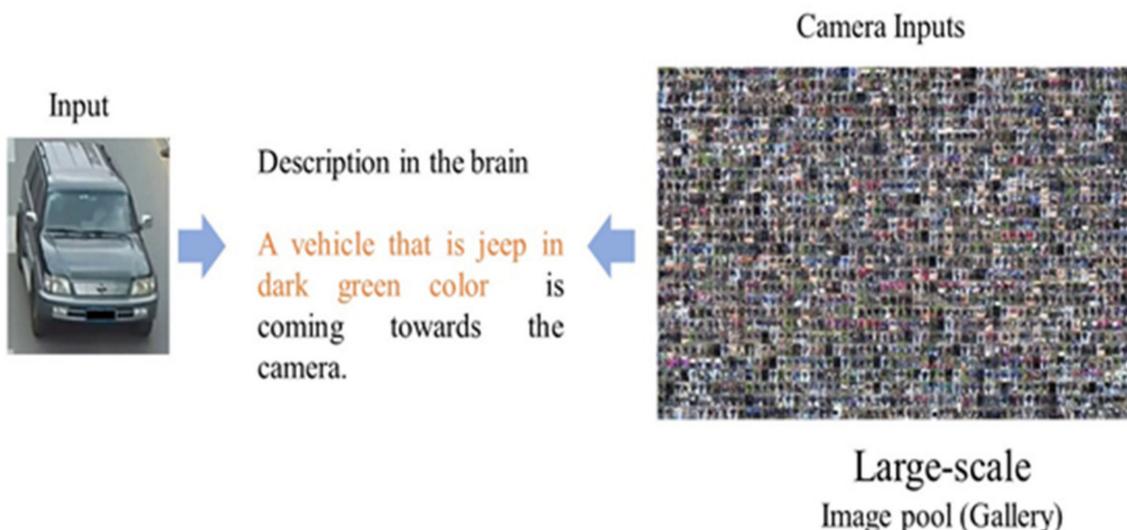
Figure 4. Shows how human re-identify vehicle?
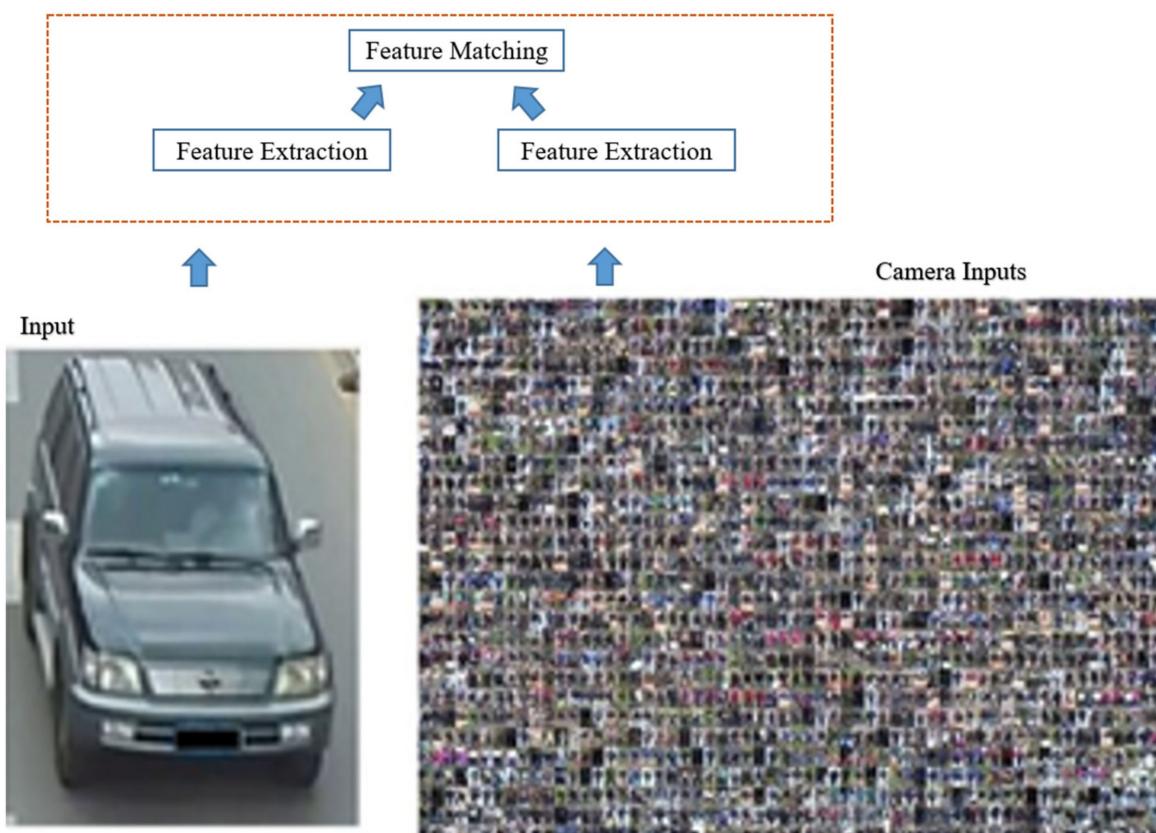
Figure 5. Illustrates how machine re-identify vehicle?

1.4. Vehicle Re-Identification Practical Application

There are many significant real-world applications where vehicle re-id system can be utilized and satisfies the great needs of our practical life. However, some major applications are briefly discussed as follows:

- Suspicious vehicle search: Most of the time terrorists use vehicle for their criminal activities and soon leave that spot on vehicles. It is very difficult to fast search suspicious vehicle manually from surveillance camera.

- Cross camera vehicle tracking: In vehicle race sports, some of the viewers on television wish to watch specific vehicle. With vehicle re-id system broadcaster can only focus on that specific vehicle when it comes in the field of view of surveillance camera network.
- Automatic toll collection: Vehicle re-id system can be used at toll gates to identify vehicle type like small medium and large and charge the toll rate accordingly. Automatic toll collection reduces delay and improves the toll collection performance by saving travelers time and fuel consumption.
- Road access restriction management: In big cities, heavy vehicles like trucks are not permitted in the daytime, or some of the vehicles with specific license plate number are permitted on specific days to avoid congestion in city or officially authorized vehicles can enter in city.
- Parking lot access: vehicle re-id system can be deployed at the gate of parking lot of different places like head offices, and residential societies. So only authorized vehicles are allowed to park.
- Traffic behavior analysis: Vehicle re-id can be used to examine the traffic pressure on different roads at different times, such as peak hours calculation or particular vehicle type behavior.
- Vehicle counting: System can be useful to count a certain type of vehicle.
- Speed restriction management system: Vehicle re-id system can be utilized to calculate the vehicle's average speed when it is crossing from two subsequent surveillance camera positions.
- Travel time estimation: Travel time information is important for a person who is traveling on road, it can be calculated when a vehicle is passing in between consecutive surveillance cameras.
- Traffic congestion estimation: By knowing the number of vehicles flow from one point to another point within a specific time period using vehicle re-id system, we can estimate traffic congestion at the common spot from where all vehicles may cross.
- Delay estimation: Specific commercial vehicle delay can be estimated after predicting traffic congestion on the rout that vehicle follows.
- Highway data collection: Highway data can be collected through surveillance cameras that are installed on roadsides and that data can be used for any purposes after processing and analyzing at the traffic control center.
- Traffic management systems (TMS): Vehicle re-id is an integral part of TMS, it helps to increase transportation performance, for instance, safe movement, flow, and economic productivity. TMS gathers the real-time data from the surveillance cameras network and streams into the Transportation Management Center (TMC) for data processing and analyzing.
- Weather precautionary measures: When specific vehicle is identified that may be affected by weather, then traffic management systems notify that vehicle about weather conditions like wind velocity, severe weather etc.
- Emergency vehicle pre-emption: If any suspicious vehicle is identified at any event or road then vehicle pre-emption system passes messages towards lifesaving agencies such as security, firefighters, ambulance, traffic police, etc. to reach in time and stabilize the scene. With this system, we can maximize safety and minimize response time.
- Access control: Vehicle re-id system can be implemented for providing safety and security, logging and event management. With the implementation of the system only authorized members can get an automatic door opening facility, which helps guards on duty.
- Border control: Vehicle re-id system can be adopted at different check posts to minimize illegal vehicle border crossing. Vehicle re-id system can provide vehicle and owner's information as it approaches security officer after identifying the vehicle. Commonly these illegal vehicles are involved in cargo smuggling.
- Traffic signal light: When the traffic light is red and any vehicle crosses stop line, the vehicle re-id system can be implemented to identify that vehicle for fine.

- Vehicle retrieval: In this case, re-id is associated with a recognition task. The specific query with a target vehicle is provided, and all the related vehicles are searched in the database. The re-id task is thus employed for image retrieval and usually provides ranked lists, similarly related items, and so on.

However, due to the vast range of practical applications that employ vehicle re-id system and to limit the scope of the paper, this review article mainly focuses only on vision-based methods. Moreover, it is very hard to cover all technologies for vehicle re-id in one survey paper but despite of that we have summarized the strengths and weaknesses of all technologies in Table 1. Therefore, this review article focuses on the use of vision-based approaches including, Appearance, license plate, contextual information etc. In last few years, there has been lack of comprehensive study of the overall problem and different solutions. This paper fills the gap by providing a detailed review covering main challenges, different approaches, and applicability. In addition, it provides the analysis and comparison of existing vehicle re-id methodologies. Aiming to facilitate other researchers, this review also provides the required information about the publicly available datasets and discusses several important research directions with under-investigated open issues to narrow the gap between the closed-world and open-world applications, taking a step towards real-world re-id system design.

Table 1. Summary of strengths and weaknesses of different vehicle re-id technologies.

Technology	Strengths	Weaknesses
Surveillance camera	Don't require the owner's cooperation. Low cost because usually cameras are installed on roadsides, so don't require additional charges to install.	Complex and unconstrained environment along with varied road topology affects the performance. Performance degrades due to dirt, snow, occluded image, blurry image, and sunshine, etc. The vehicle is identified only when it comes in the field of view of the camera.
Magnetic Sensor	Insensitive to bad weather like snow, fog, and rain. There is no privacy issue in magnetic sensors.	Complicated installation. Embedding magnetic sensor under carriageway after drilling hole. Identified only at the detection terminal.
Inductive loop	Provides different traffic parameters like speed, volume, headway, presence, and occupancy etc.	Installation of inductive loop technology requires metallic loops under the road. The vehicle is identified in the field-of-view of detection terminal.
GPS	Provides continuous vehicle information, such as space and time, to the control centre. 100% vehicle recognition rate.	Require owner's cooperation to install hardware in vehicle. Varying accuracies, minimal fleet penetration, and signal loss because of tunnels, trees, and tall buildings.

Two ways for writing surveys can be found in the object re-id literature; first way gives a deep insight into methodologies, whereas the second way covers the overall perspective related to the problem [17,18] This survey includes both methodologies and overall perspective of vehicle re-id literature. We also review the recent development of vision-based vehicle re-id along with other technologies. In addition, this survey draws a timeline to introduce important milestones for vehicle re-id, which can be seen in Figure 6.
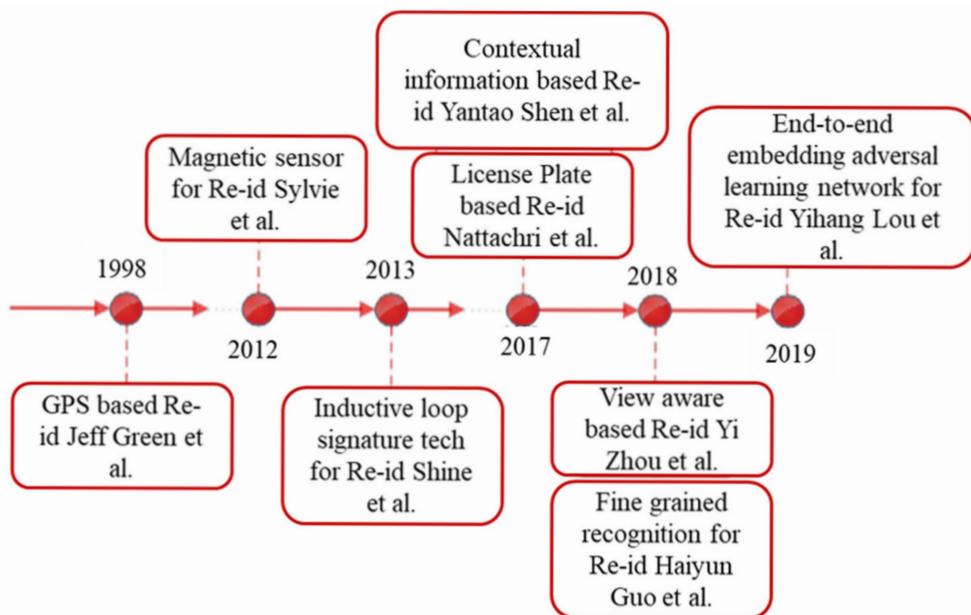
Figure 6. Milestones existing re-id approaches in the Vehicle re-id history.

The paper is organized in the following way. Sections 2–5 provide an overview of recent state-of-the-art proposed methodologies in various technologies. Section 6 presents a publicly available benchmark dataset that covers various real-world surveillance scenarios. Section 7 discusses the challenging problems in vehicle re-id. Section 8 sheds light on the evaluation measures for vehicle re-id. Section 9 analyzes and compares the experimental results of various approaches. Meanwhile, the last section concludes and discusses future work.

The main contributions of this review paper is summarized as follows:

- To the best of our knowledge, this is the first comprehensive review paper that covers computer vision-based methods for vehicle re-id tasks, with a different technological background of approaches for completeness such as, global positioning systems (GPS), inductive loop and magnetic sensors.
- Discusses various real-world applications of vehicle re-id in different domains including the intelligent transportation system.
- Comprehensive comparisons of existing methods on several state-of-the-art publicly available vehicle re-id datasets are provided, with brief summaries and insightful discussions being presented.
- Discusses the challenges in detail for designing an efficient vehicle re-id system and illustrates the recent trends and future directions.

2. Methods Used for Vehicle Re-Identification

Traditionally different traffic sensors are adopted to know the vehicle presence, volume, occupancy, and speed data. Nowadays, new sensor-based technology is adopted to get more information like origin-destination estimation, travel time and other travel information applications. Based on different technologies vehicle re-id approaches can be divided into six categories, as depicted in Figure 7.
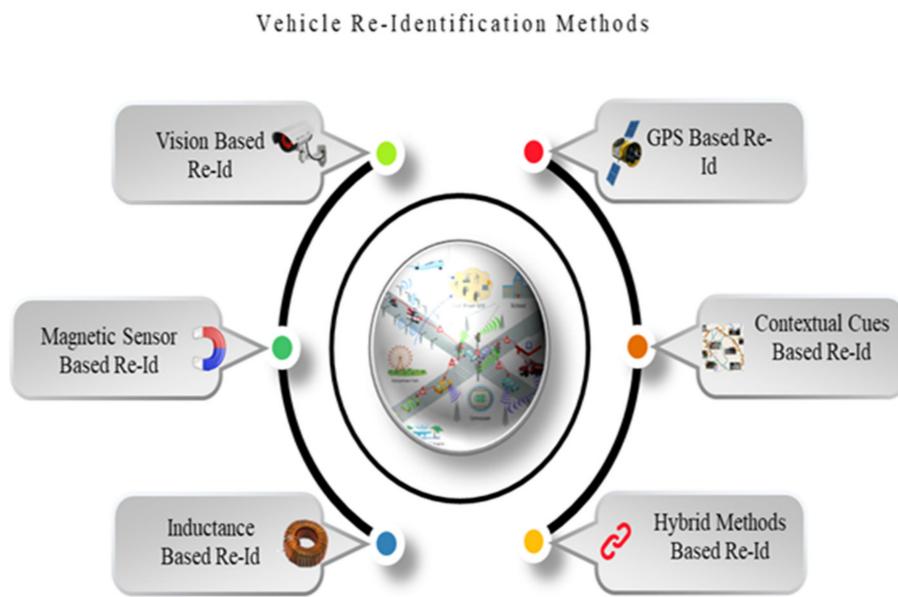
Figure 7. Shows vehicle re-id methods.

2.1. Magnetic Sensor-Based Vehicle Re-Identification

An electromagnetic field is used to detect the vehicle, when it crosses and it is used to provide occupancies, counts, and vehicle speed. However, vehicles are made up of metal. It disrupts the magnetic field, so magnetic signature regenerated by one vehicle is different from the other vehicle [19]. This approach helps in re-identifying a specific vehicle. Moreover, for ITS the Berkeley's company sells magnetic sensors with the name "Sensys Network" [20]. A straight-line re-id rat is 50%, and the approach reduces the magnetic signature peak value sequence for calculating the signature distance to prevent vehicle speed dependency [21]. For real-time vehicle re-id processing unit is associated to thousands of magnetic sensor nodes and a large number of magnetic sensors that generate massive data streams, and to deal with real-time data stream mining, high-performance FPGAs and low-performance microcontroller are used [22,23]. Sylvie Charbonnier et al. [24] studied various approaches for vehicle re-id by adopting vehicle tridimensional magnetic signature measured with sensor, when car passes sensor and changes in the magnetic field were induced and measured in three different directions like X, Y, Z. Rene O. Sanchez et al. [25] investigated vehicle re-id approaches by using wireless magnetic sensors and compares vehicle magnetic signatures to overcome the limitations of system while vehicle is stopped or moving slow at detection station.

2.2. Inductive Loop-Based Vehicle Re-Identification

Vehicle can be re-identified using inductive loops embedded in the road surface for the detection of vehicle. From those loops, a fingerprint is captured for every car passing by. The travel time can be determined when those fingerprints or certain aspects of them coming from different locations are compared with each other. Jeng and Chu [26] designed a real-time inductive loop signature-based vehicle re-id method named RTREID-2M. Inductive signature is used for vehicle re-id and much efforts have been done to utilize inductive loop signature technology. Inductive signature-based vehicle re-id algorithms identify specific vehicle at downstream detection station by matching the inductive signature at upstream detection station, considering that vehicle have same signature by crossing different loop detection stations [27]. Vehicle re-id researchers have proposed several algorithms like optimization, piecewise slope rate (PSR) matching [28], lexicographic and blind deconvolution [29], all these proposed approaches are for raw signature processing, signature

feature extraction, and vehicle matching. R.J. Blokpoel [30] proposed an algorithm with different sizes of a single loop. Validation tests depict re-id rates up to 100%, when loops are identical to the similar type and 88% when compare between different types.

2.3. Global Positioning Systems-Based Vehicle Re-Identification

Global Positioning Systems (GPS) technology is an essential and valuable tool for ITS and traffic surveillance, because it provides positioning data for every single vehicle [31,32]. There are still some limitations in vehicle re-id using GPS like varying accuracy, minimal fleet penetration, and signal loss because of tunnels, trees, tall buildings, etc. GPS is adopted with vehicles to locate and get travel information along with longitude and latitude information and timestamp. GPS is special form of mobile sensing technology that enables the devices like GPS logger, GPS cellular phones, and smartphones moves with vehicles to get speed information and location continuously. However, different types of vehicles have different behaviors such as deceleration rates, acceleration, and speed variation. This encourages the author to adopt GPS technology for vehicle classification and re-id [33].

2.4. Vision-Based Vehicle Re-Identification

In computer vision, the aim of vehicle re-id is to identify specific vehicle that appeared over in multiple cameras network. The large surveillance camera network is deployed in different areas of public places like hospitals, parks, colleges, roads, and other areas. It is also difficult and tiresome job for security officers to track targeted or specific vehicle over multiple camera network manually. However, computer vision techniques can automatically re-id a vehicle and basic five main working steps are discussed below (shown in Figure 8).
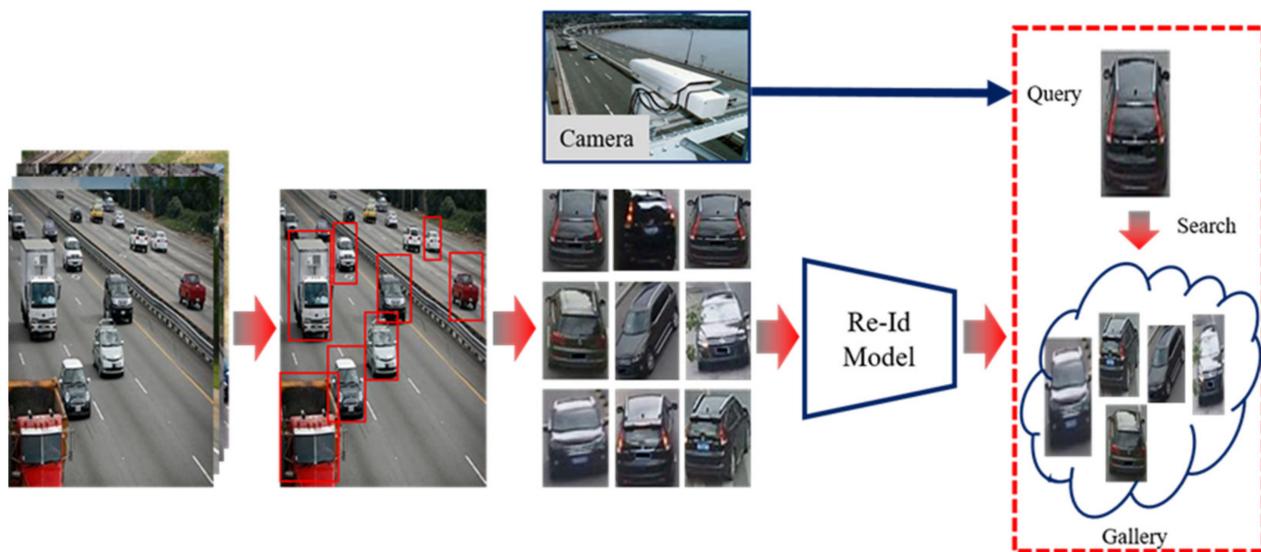
Figure 8. The flow of designing a practical vehicle re-id system, including five main steps.

- Step 1: Data Collection: For real-time video analysis raw videos from surveillance cameras is one of the key component. The cameras are fixed at different locations in an unconstrained environment [34].
- Step 2: Bounding Box Generation: It is very difficult almost impossible when we have large scale surveillance videos to extract vehicle image. We use a bounding box and it is obtained by vehicle detection technique [35].
- Step 3: Training Data Annotation: Data annotation is a process of labeling the videos or images of dataset with metadata. It is an indispensable step for vehicle re-id model training because each surveillance camera video recording is in a different environment.
- Step 4: Model Training: Model training is simply the task of learning discriminative features and good values for all the weights and the bias from previous annotated

vehicle videos or images of the dataset. It is a key step in vehicle re-id systems and a widely explored area in literature.

- Step 5: Vehicle Retrieval: Vehicle retrieval is a task of matching targeted vehicle (query image) over a gallery set.

3. Vision-Based State-of-the-Art Vehicle Re-Identification Approaches

Vision-based methods focus on examining robust feature representations to calculate the distance between features of two-vehicle images and vehicles with the same class have a low distance otherwise high. However, vehicle features are difficult to distinguish when a captured vehicle image consists of similar colors and pose. In this section gives an overview of recent works on computer vision-based methods for vehicle re-id problem, furthermore general approach for vision-based method is shown in Figure 9. Several impressive vision-based methods have been proposed to improve vehicle re-id performance either by modifying the existing DL architectures or designing a new deep neural network (DNN). Generally speaking, eight different techniques have been employed in this research area: (A) Feature representation for vehicle re-id, (B) Similarity metric for vehicle re-id, (C) Traditional machine learning-based vehicle re-id, (D) View-aware-based vehicle re-id, (E) Fine-grained visual recognition-based vehicle re-id, (F) Generative adversarial network-based vehicle re-id, (G) Attention mechanism, (H) License plate-based vehicle re-id.
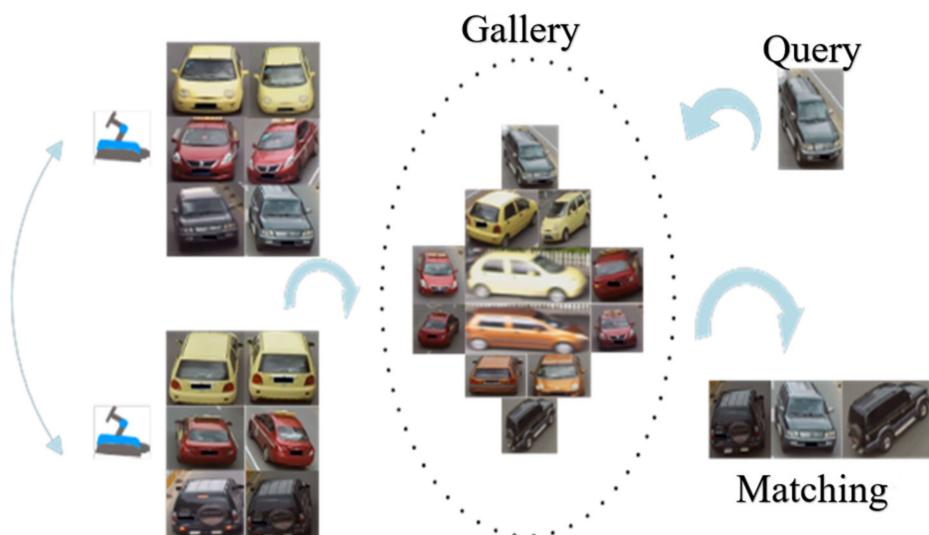
Figure 9. The vehicle re-id problem: given a Query, find the matching candidate in the gallery.

3.1. Feature Representation for Vehicle Re-Identification

Feature representation play vital role in progress of many different computer vision tasks. In this regard, vehicle re-id features representation approaches can primarily be classified into two parts: hand-crafted and deep learning features representations. Hand-crafted feature representations BOWCN [36], and LOMO [37] initially utilized in person re-id and then applied directly on vehicle re-id task. Some well-known deep learning-based feature representations such as GoogLeNet [38], VGGNet [39], AlexNet [40], and, ResNet [41] are used for vehicle re-id. The researcher also adopts these baseline models in their approaches for vehicle re-id. Such as, NuFACT [42] takes GoogLeNet [38], FACT [43] uses AlexNet [40], DRDL [44] utilizes VGGNet [39] to extract features of vehicles. Various type of loss functions are utilized to efficiently learn vehicle image discriminative feature representation to train deep learning-based model vehicle re-id; such as the deep joint discriminative learning (DJDL) [45] approach uses identification, and verification and triplet loss functions improved triplet convolutional neural network [46] uses classification and-oriented and triplet loss function to extract discriminative feature representation.

3.2. Traditional Machine Learning-Based Vehicle Re-Identification

In traditional machine learning (TML), we adopt feature engineering to artificially clean and refine data. However, previously proposed approaches are grouped into for robust features extraction and learning discriminative classifiers. In TML extracted features are directly computed from image pixels and it is low level feature representation. Moreover, TML-based algorithm design is expensive and difficult. Broadly, it consists of two steps feature extraction and feature classification. There are many algorithms proposed for low level feature extraction for instance speeded up robust features (SURF) [47], scale-invariant feature transform (SIFT) [48], and histogram of oriented gradient (HOG). After feature extraction different classifiers are applied, which are widely used in TML approaches such as linear regression, k-Nearest Neighbor (KNN) [49], logistic regression, support vector machine (SVM) [50], bayes classification [51], and decision tree [52]. The features extracted using SIFT are local features of the image, which maintains the scale scaling, invariance of rotation, and brightness variation. In addition, it also maintains a particular degree of stability to affine transformation, the viewing angle change, and noise.

Moreover, one of the feature descriptor adopted for targeted object detection in image processing is HOG. The large area of image features are formed by calculating the gradient direction histograms of its local regions. However, an overlapping local contrast normalization approach is adopted to improve the performance. Zapletal and Herout [53] utilize the color histogram and the HOG features with linear regression to re-id vehicle. Chen et al. [54] designed a method to re-id vehicles grid-by-grid with HOG features extraction for coarse search and further improves the result by utilizing histograms of matching pairs. In [55], vehicle re-id local variance measures are applied using local binary patterns and joint descriptors.

3.3. Similarity Metric for Vehicle Re-Identification

Performance of vehicle re-id can be improved by selecting appropriate distance matrices regardless of appearance representation. Distance metric learning approaches [56] are thoroughly studied in image retrieval and recognition tasks, in which matric space is defined in such a way that features that belong to same class are kept closer and different are at distant as shown in Figure 10. In the re-id task, image features are known as appearance descriptor. In this the learned distance matric in appearance space minimizes the distance for descriptor between same vehicles and maximizes distance for descriptor of different vehicles. As in various face recognition algorithms [57,58] uses Euclidean and Cosine distance matric to measure the similarity, and FACT [43] also utilizes Euclidean and cosine distance metrics to measure similarity between the pair of vehicle for re-id. Similarly, NuFact [42] utilizes the Euclidean distance to measure the similarity between the probe and gallery set vehicle images in discriminative null space [59]. Furthermore, deep relative distance learning (DRDL) [44] studied a two-branch convolutional neural network to covert the raw vehicle images into a Euclidean space, so that distance can be used directly to measure the similarity of two individual vehicles.

Pairwise constraints are required for matrix learning and it is done in supervised fashion. During the training features of appearance descriptor are in pair and labelled as positive and negative. It is totally depending on appearance descriptor whether it belongs to the same vehicle or different vehicle. Appearance descriptors are represented as x_1, x_2, ..., x_n, here n represents number of training instances and the dimensionality of every instance is represented by m. The aim of metrics learning is to learn distance metric and matrix $D \in \mathrm{R_{mxm}}$ represents it; thus, the distance between pair of appearance descriptors x_i and x_j is as follows:

$$d\left(x_i, x_j\right) = \left(x_i - x_j\right)^T D\left(x_i - x_j\right) \tag{2}$$

$d(x_i, x_j)$ is a true metric only possible when matrix D is symmetric positive semi-definite. This issue is resolved by adopting convex programming as follows:

$$min_D \sum_{(x_i, x_j) \in Pos} \|(x_i - x_j)\|_D^2 \ s.t. D \geq 0, \ and \sum_{(x_i, x_j) \in Neg} \|(x_i - x_j)\|_D^2 \geq 1 \qquad (3)$$

where *Pos* represents the positive label in training samples, and it is the appearance descriptor of the same vehicle, whereas *Neg* represents the negative label in training samples and it is the appearance descriptor of a different vehicle.
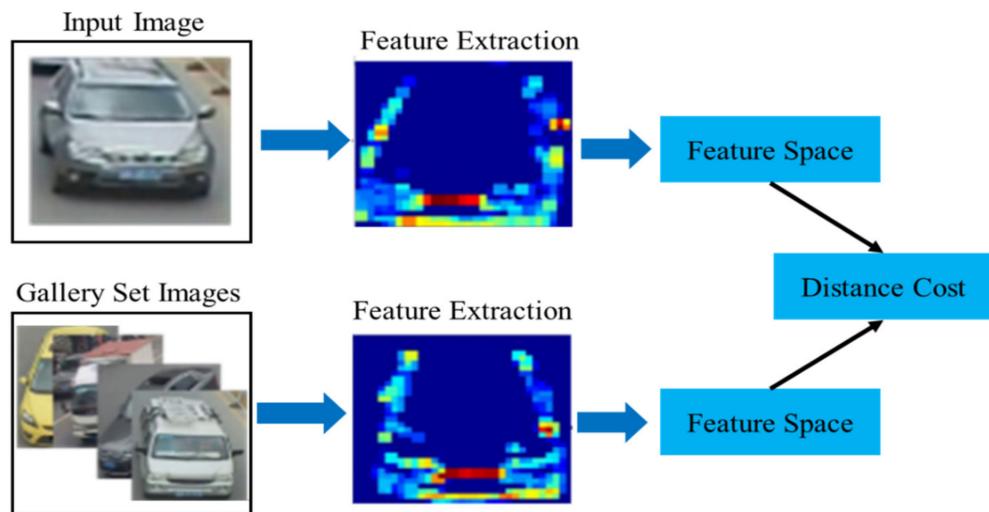
Figure 10. Vehicle re-id system based on metric-based methods.

3.4. Fine-Grained Visual Recognition-Based Vehicle Re-Identification

Vehicle re-id is fine-grained recognition task, and fine-grained vehicle recognition can be divided into two parts, representation learning model and part-based model. Many approaches are proposed [60] that utilize alignment and part localization for feature extraction of main parts and then those parts are compared for vehicle re-id. Xiao et al. [61] studied weakly supervised way in fine-grained domain using reinforcement learning to get discriminative parts of vehicle. In addition, Lin et al. [62] presents a bilinear architecture to get the pair of local features in which output descriptors of two networks are merged in an invariant way. Boonsim et al. [63] presents an approach for fine-grained recognition of vehicles at night. The authors utilize shape and lights of vehicle visible in night and relative position to identify model and make of a vehicle, which are visible from the front and rear side.

In fine-grained recognition, local region features are extracted from different points such as logo, annual inspection stickers, and decorations, to make system more efficient and robust various attributes of vehicles are also incorporated like color, model, and type information. For example, in different vehicles with similar global appearance in Figure 11, all the vehicles are different in each column. The differences between each vehicle are pointed out with red circles. From Figure 11 it can also be seen that the differences between similar global appearance vehicles lie in some local regions.

Figure 11. Shows vehicles that are same in global appearance but differentiated by local regions that are marked in red circle.

3.5. View-Aware-Based Vehicle Re-Identification

Most of the above discussed deep learning features [38,39,45] are general, and these learned features end at multiple fully connected layers. Despite that, all these approaches performance is not bad. But these approaches are not designed for a specific problem related to view point variation. It is a central challenge in vehicle re-id task. Vehicle re-id is closely related to person re-id, however, intra-class variation is a major problem in person re-id in which the same person looks different by changing viewpoint. Zhao et al. [64] designed a novel approach that achieved satisfactory results and the method was based on person body parts guided for re-id. Wu et al. [65] proposed a study with pose prior that made identification efficient and robust to viewpoint. Zheng et al. [66] proposed the pose box structure that generates the pose estimation after affine transformations. It is also challenging and crucial in vehicle re-id, because image viewpoint is the same as a consequence of vehicle rigid motion. Wang et al. [67] studied the orientation invariant feature embedding to solve the issue of viewpoint variation influence on vehicle re-id system. Prokaj et al. [68] proposed a pose estimation-based approach to handle multiple viewpoint problem. Yi Zhou et al. [69] studied uncertainty in the viewpoint of vehicle re-id system and designed end to end deep learning-based architecture on Long Short-Term Memory (LSTM) bi-directional loop and concatenated CNN, in this model author takes full advantage of LSTM and CNN to learn the different viewpoints of vehicle. And also, there are many more approaches are proposed to handle the view point variation issue in vehicle re-id such as adversarial bi-directional long short-term memory (LSTM) network (ABLN) [70], spatially concatenated convolutional network (SCCN) and CNN-LSTM bi-directional loop (CLBL) [69]. However, all these approaches need vehicle datasets. Every vehicle image is densely sampled camera viewpoints. Despite that, it is hard to gain in real-time camera surveillance systems. Therefore, there is still ample room for vehicle re-id by thoroughly considering viewpoint variations.

3.6. Generative Adversarial Network-Based Vehicle Re-Identification

GAN [71] is one of the hot technique in semi-supervised and unsupervised learning algorithms. It is proposed by Goodfellow by deriving backpropagation signals through a competitive process involving a pair of networks. GAN can be adopted in different applications, like style transfer, image synthesis, image super-resolution, semantic image editing, image super-resolution, classification and person/vehicle re-id. The GAN-based vehicle re-id flow is shown in Figure 12. At present, there have been many papers that adopt GAN to solve the problems of vehicle re-id. The existing datasets have low diversities and

small scales, which leads to poor generalization performance on the trained models. To solve this problem. Generative Adversarial Network (GAN) in Object re-id is among the latest research trends in the deep learning approaches. GANs achieved significant performance in in many fields such as translation [72] and image generation [73]. Furthermore, recently GANs are also utilized for re-id problems (person re-id and vehicle re-id) [74,75]. Zheng et al. [76] proposed a method in which they used the DCGAN [73] with Gaussian noises to generate unlabeled person images before training. Wei et al. [77] studied a PT-GAN to minimize the domain gap by transferring person images between different styles. Zhou et al. [78] proposed GAN based model to solve cross-view vehicle re-id problem by generating vehicle images in different viewpoints. Lou et al. [74] designed a model to generate the same and cross-view vehicle images from original images to facilitate training model. Zhou et al. [78] proposed a conditional generative network to generate cross-view images from desired vehicle pairs.

Figure 12. Vehicle re-id system based on GAN diagram.

Aihua et al. [79] proposed a framework that primarily comprises view transform and vehicle re-id model. The view transform model comprises of GAN to generate vehicle images in different views to overcome the viewpoint related issue. The vehicle re-id model consists of one backbone, three subnetworks, and one embedding network. The overall framework is illustrated in Figure 13.
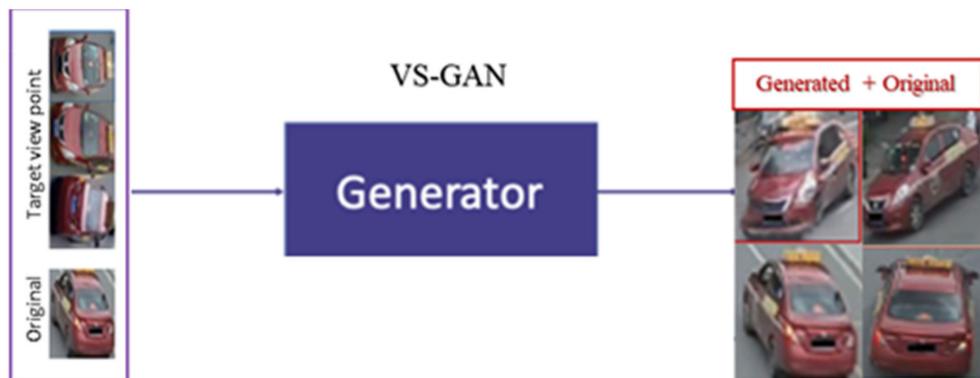
Figure 13. Overview of deep feature representations guided by the meaningful attributes.

3.7. Attention Mechanism

The neural networks at some extent imitate human brain actions in simple way. Attention Mechanism is also an effort to develop a technique that concentrate on selective thing/actions that are relevant to task and neglecting the others in neural networks. Currently, researchers are trying hard to design an efficient attention-based neural network for vision-related applications. Such as image classification [80], fine-grained image recognition [81], action recognition [82], and re-id [83]. The commonly followed strategy in these approaches is integrating a hard part selection subnet work or soft mask branch into the deep networks. Such as Zhao et al. [84] studied the part-localization CNN for predicting salient parts and features of these parts exploit for person re-id. Wang et al. [80] utilizes residual learning technique [41] to develop the Residual Attention unit for soft mask learning and gained significant image classification results. Though, only the soft pixel-level attention has very small participation in the performance of vehicle re-id task. It gives only global information like vehicle logo, annual inspection stickers, and personalized decorations. So, they presented joint learning framework for vehicle re-id in which both soft and hard level attentions are utilized Furthermore, Guo et al. [85] proposed a model with one

trunk and two salient part branches for hard part level attention. Trunk branches extracts the global features of vehicle and salient branches extracts the features from vehicle head parts and windscreen. For soft pixel level attention residual attention modules are inserted into trunk and salient branches. Lastly, global and salient part features of vehicle are put to gather for effective feature representation with the supervision of multi-grain ranking loss for vehicle re-id task and complete framework is shown in Figure 14. Furthermore, comparison of different attention mechanism-based approaches are shown in Table 2.
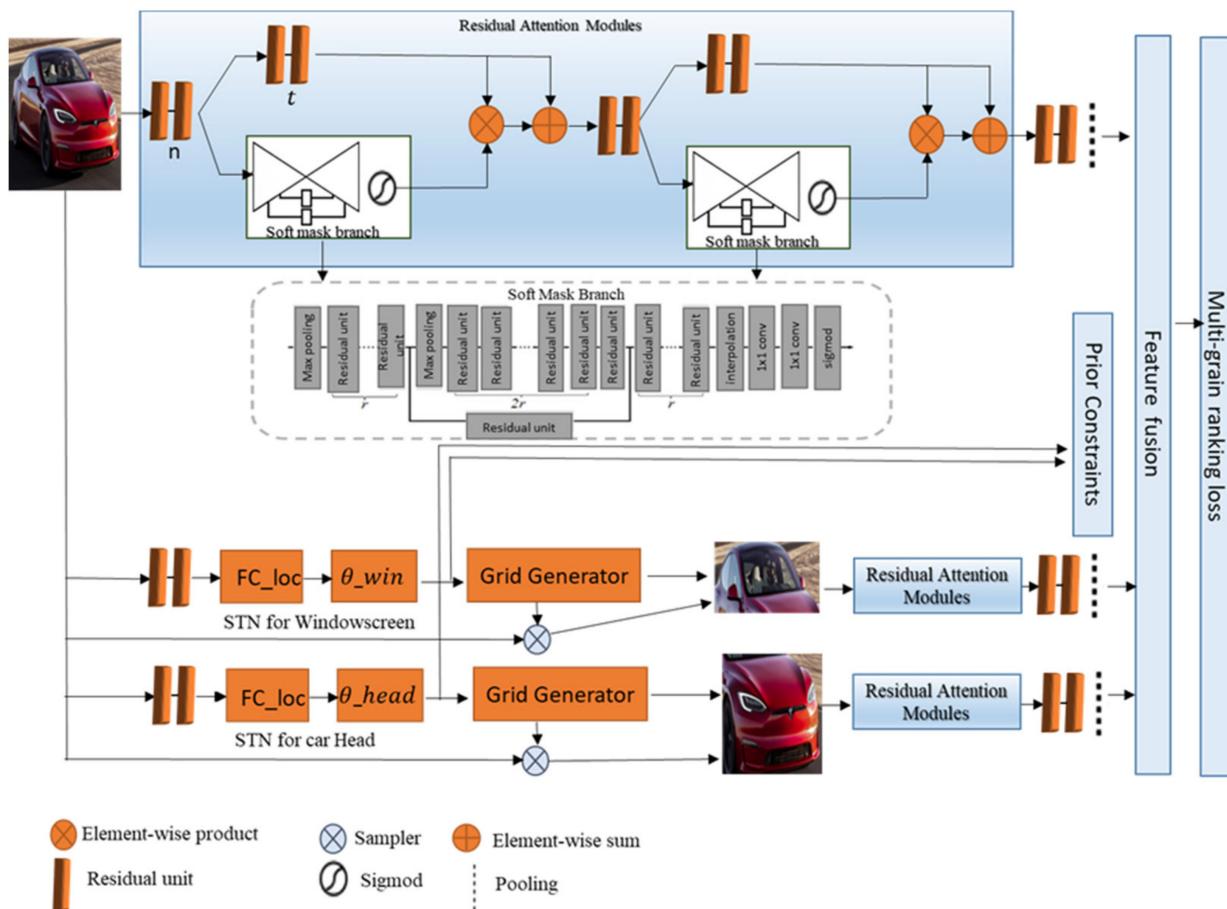

Figure 14. An overview of Two-level Attention network supervised by a Multi-grain Ranking loss (TAMR) structure.

Table 2. Comparison of different attention mechanism-based approaches.

Method and Reference	mAP%	Rank-1 (%)	Rank-5 (%)
RNN-HA [86]	56.80	74.79	87.31
SCAN [87]	49.87	82.24	90.76
AAVER [88]	61.18	88.97	94.70
PGAN [89]	79.30	96.50	98.30

3.8. License Plate-Based Vehicle Re-Identification

Vehicle re-id using license plate is simply the system's ability to automatically detect, extract, and recognize license plate characters automatically from vehicle image. License plate recognition (LPR) is a conventional method to identify a specific vehicle [90]. An automatic LPR system is mainly divided into two parts, first license plate detection and second, interpreting the vehicle license plate image into numerically readable form. There are many approaches proposed in past for LPR. However, it is still challenging due to some reasons like vehicle image is not captured perfectly, some characters may be occluded, illumination, variation in size of an image, camera distance and zooming. Li and Shen [91]

studied a sequence labelling-based approach to recognize the vehicle license plate without character-level segmentation using recurrent neural networks (RNN). The input feature sequence to RNN is extracted using a nine-layer CNN. Super-resolution is also proposed to restore a license plate image to improve performance. Shi et al. [92] designed convolutional recurrent neural network (CRNN) for scene text recognition that incorporates feature extraction, transcription and sequence modeling into a unified framework. Moreover, Figure 15 shows the basic steps of license plate-based vehicle re-id.
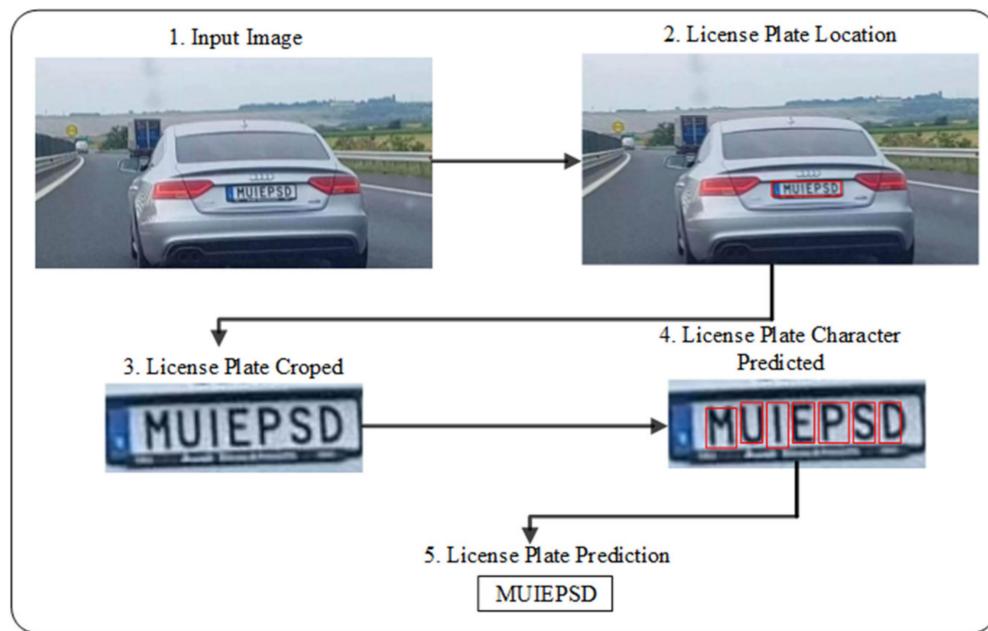
Figure 15. The flow of the license plate-based vehicle recognition.

4. Spatio-Temporal Cues-Based Vehicle Re-Identification Approaches

Introducing contextual information in vehicle re-id system can increase the efficiency and reduces irrelevant vehicle gallery images. As compared to person, for vehicle it is necessary to follow traffic rules for instance, practically vehicle follows speed limits, routes, and traffic lanes, so in this scenario vehicle moving in between different cameras at specific time and location helps a lot in vehicle re-id. Spatio-temporal cues are greatly examined for various objects association in surveillance camera network [93]. As in [94] concluded few key findings. Firstly, one specific captured vehicle in one camera cannot appear at more than one location at the same time. Secondly, along the time vehicle is moving continuously based on these finds, authors use location and time slots to eliminate irrelevant vehicle images from list as demonstrated in Figure 16. Ellis et al. [93] proposed approach that trains the model on temporal and topological transitions of trajectory data and is acquired from surveillance camera network. Loy et al. [95] presented a method for obtaining the spatio-temporal topology of surveillance camera network using multiple camera correlation analysis. Furthermore, time and location information is also exploited for vehicle re-id task. Liu et al. [96] studied a spatio-temporal affinity method for quantifying different pairs of vehicle images. Shen et al. [97] also introduces the spatio-temporal path data for vehicle re-id.

Figure 16. Depicts the spatio-temporal information.

5. Hybrid Methods-Based Vehicle Re-Identification

To further enhance the robustness and efficiency of vehicle re-id system researchers have proposed the approaches in which they combined the two or more different techniques, for instance Liu et al. [42] proposed a framework with name PROVID, in this framework author not only consider the visual appearance of vehicle for re-id system, but also exploits the license plate and spatio-temporal cues of vehicle as shown in Figure 17. Jiang et al. [98] studied vehicle re-id algorithm using appearance and contextual information, author examines the multiple attributes during training like vehicle model, color, and vehicle image features individual respectively and sort vehicles on the bases of spatio-temporal cues. Shen et al. [97] designed a two-step architecture, a pair of query vehicle images with contextual information and visual temporal path are produced using Markov Random Fields (MRF) chain model, and then the similarity score is generated.
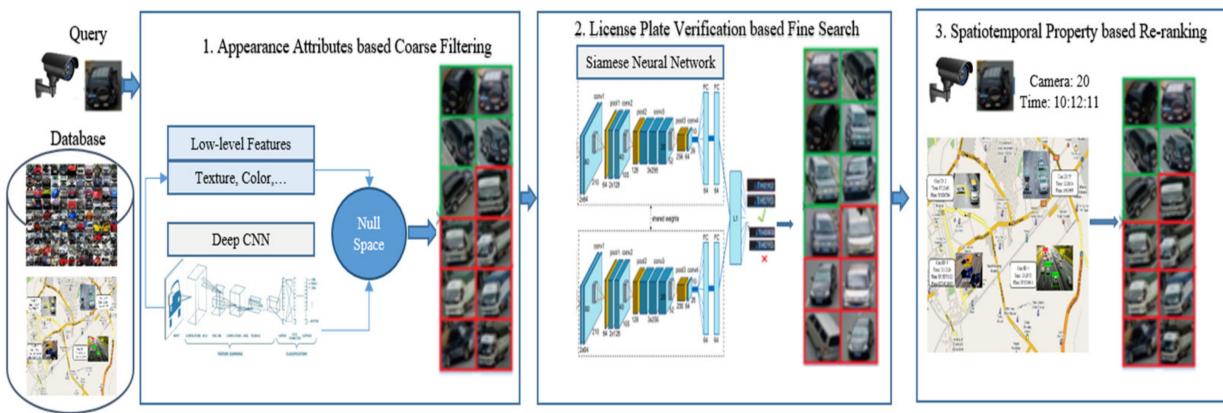
Figure 17. The architecture of the PROVID framework.

## 6. Vehicle Re-Identification Benchmark Datasets

Datasets are the key components to measure the performance of vehicle re-id system and should reflect the practical surveillance camera data. We cannot avoid some factors like occlusion, background clutter, change in illumination etc. to evaluate the approach [99]. However, multiple benchmark datasets are available, some well-known like VeRi-776,

VehicleID, etc. that are prepared by the research community to evaluate vehicle re-id techniques. Table 3 and Figure 18 lists the commonly used vehicle re-id dataset with attributes. Furthermore, a brief description of the most popular datasets is as follows:

Table 3. Characteristics of publicly available datasets.

S. No	Dataset	Year	Total No. of Images	No. of Vehicle Models	No. of Vehicles	No. of Viewpoints	No. of Cameras
1	VeRi-776 [43]	2016	50,000	10	776	6	18
2	PKU VehicleID [44]	2016	221,763	250	26,267	2	12
3	Vehicle-1M [100]	2018	936,051	400	55,527	......	......
4	BoxCars21k [35]	2016	63,750	148	21,250	4	......
5	VehicleReId [53]	2016	47,123	......	1232	......	......
6	CompCars [101]	2015	136,726	1716	......	5	......
7	VRIC [102]	2018	60,430	......	5622	......	60
8	VRID [103]	2017	10,000	10	1000	......	326
9	VERIWild [104]	2019	416,314	......	40,671	Unconstrained	174
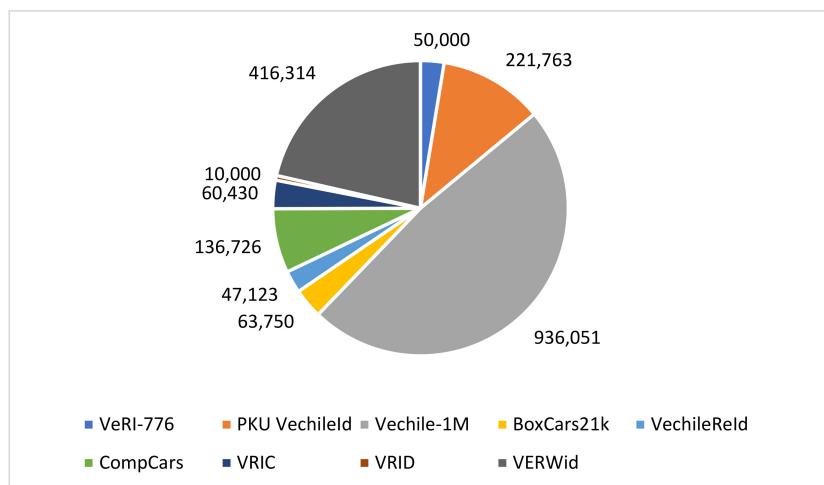
Figure 18. Depicts the number of total images per vehicle re-id dataset.

VeRi-776: [43] VeRi-776 is a publicly available vehicle re-id dataset, and often adopted by the computer vision researcher community. Dataset images are gathered in real scenario using surveillance cameras, and the total images in dataset are 50,000 of 776 different vehicles. Each captured vehicle images have 2 to 18 viewpoints with different resolution, occlusion, and illumination. Furthermore, spatio-temporal relations and license plate are annotated for all vehicles. To make dataset more robust, images are labelled with color, type, and vehicle model. In Figure 19 various types of vehicles from VeRi dataset are shown.

Figure 19. Depicts the sample images of VeRi-776 dataset.

PKU VehicleID: [44] VehicleID dataset is developed by Peking University with the funding of the Chinese national natural science foundation and national basic research program of China in the national engineering laboratory for video technology (NELVT). The vehicle dataset consists of 221,763 total images of 26,267 vehicles, and all the images are captured during daytime in a small town of China with multiple surveillance cameras with 10,319 vehicles model information i.e "Audi A6L", "MINI-cooper" and "BMW 1 Series" are labeled manually. In Figure 20 different vehicles from PKU vehicleID dataset are shown.

Figure 20. Depicts the sample images of PKU VehicleID dataset.

Vehicle-1M: [100] Vehicle-1M dataset is developed by the University of Chinese Academy of Sciences in the National laboratory of pattern recognition, Institute of Automation. This benchmark dataset contains 55,527 vehicles with 400 different vehicle models, and the total captured images are 936,051. Surveillance cameras capture all the images in China's town at day and night time and consist of a vehicle's rear and head view. Moreover, each image in this dataset is labeled with a model, make, and vehicle year. Images from Vehicle-1M are shown in Figure 21.

Figure 21. Depicts the sample images of vehicle-1M dataset.

BoxCars21k: [35] BoxCar116k dataset is developed using 37 surveillance cameras, and this dataset consists of total images 116,286 of 27,496 vehicles. For the preparation of dataset, 45 brands of the vehicle are used. Moreover, captured images of the vehicle in the

dataset are in an arbitrary viewpoint, i.e., side, back, front, and roof. All vehicle images in the dataset are annotated with 3D bounding box, model make, and type. However, some sample images are shown in Figure 22.

Figure 22. Depicts the sample images of BoxCar21k dataset.

VehicleReId: [53] VehicleReId dataset provides 47,123 vehicle images and all these images are extracted from five different video shots by using two surveillance cameras, out of total images 24,530 vehicle image pairs are human annotated.

CompCars: [101] CompCars dataset consists of two types of image nature (1) Web-nature images (2) Surveillance-nature images. There are total of 136,726 web-nature images in which there are 163 car makers with 1716 car models. However, in surveillance-nature, the total car images are 50,000 that are captured from the front view. Samples of CompCars dataset are shown in Figure 23.
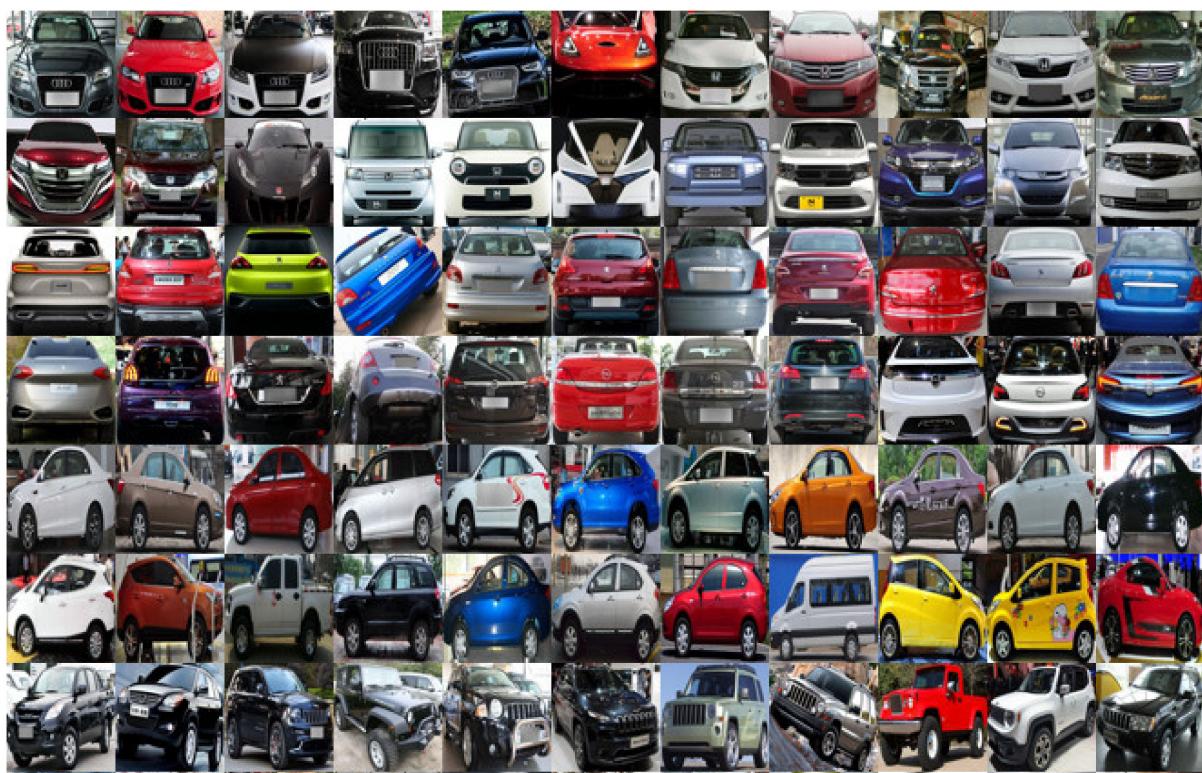
Figure 23. Depicts the sample images of CompCars dataset.

VRIC: [102] VRIC contains 5622 vehicles with 60,430 total images with different traffic road surveillance cameras and images captured at day and night. Images with different angles, viewpoints, occlusions and illuminations from VRIC dataset are depicted in Figure 24.
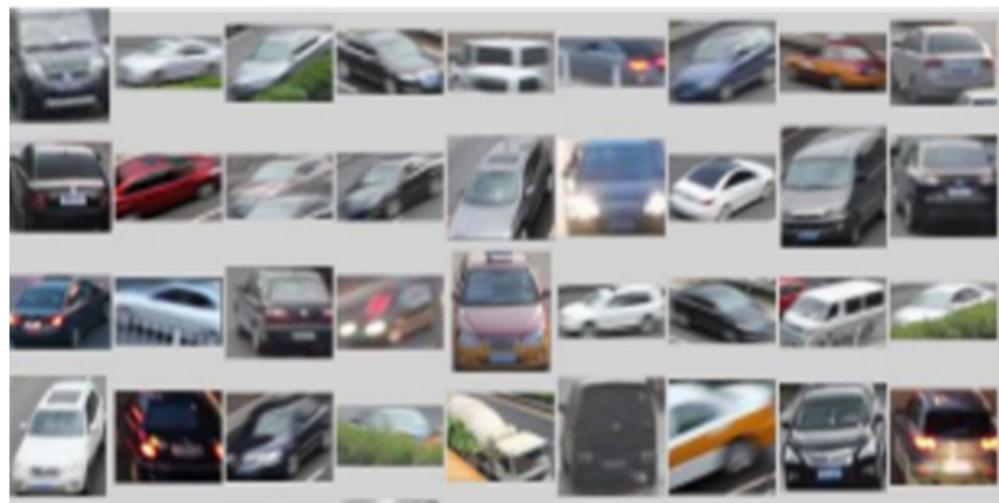

Figure 24. Depicts the sample images of VRIC dataset.

VRID: [103] This dataset contains total 10,000 images and specially developed for vehicle re-id with 326 surveillance cameras the VRID images were captured from 7 a.m. to 5 p.m. for one week. In the development of the dataset there are 1000 vehicles used with 10 commonly used vehicle models, and at least 10 times each vehicle is captured over a camera network in Guangdong city, China. Surveillance cameras have been fixed in a practical environment with arbitrary directions and angles; therefore, dataset images have various resolutions and poses distributed from 400×424 to 990×1134 pixels.

VERI-Wild: [104] Collects a large-scale vehicle re-id dataset in the unconstrained environment. For dataset development, an existing large CCTV system is utilized. It consists of 174 cameras across, recorded till one month (30×24 h). The CCTV cameras are spread over a large city consists of 200 km^2. The dataset includes 12 million vehicle raw images, and 11 volunteers cleaned the dataset for one month. After data cleaning and annotation, 416,314 vehicle images of 40,671 identities are collected. VERI-Wild dataset images with viewpoint changes, illumination variations, occlusion, and background variations are presented in Figure 25, and statistics are shown in Figure 26.

Figure 25. Depicts the sample images of VERI-Wild dataset.
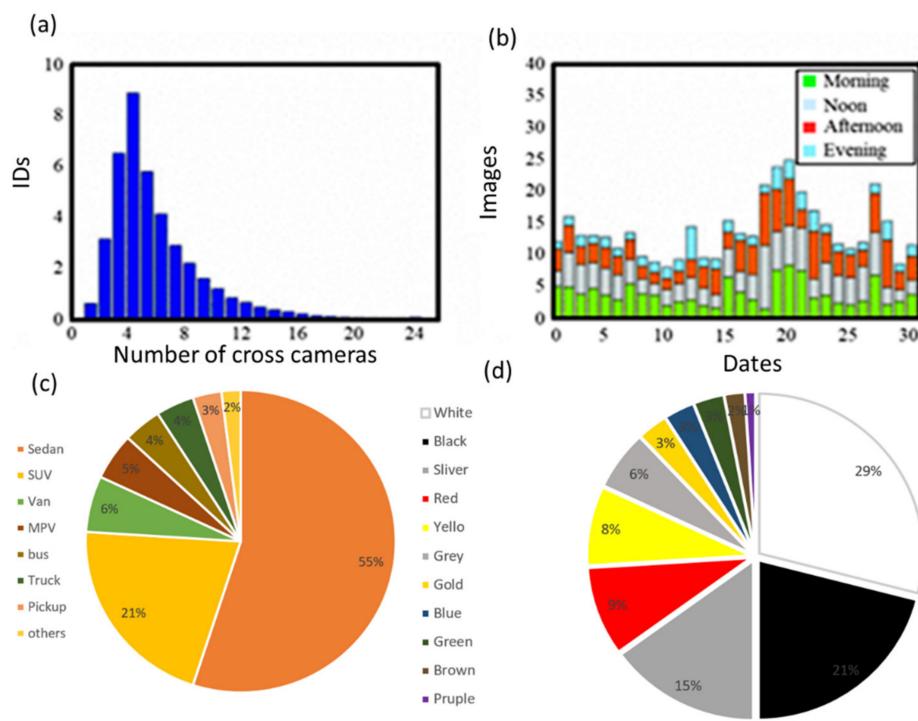
Figure 26. Illustrates the characteristics of VERI-Wild dataset. (**a**) The number of identities across multiple surveillance cameras; (**b**) Total number of IDs captured in different slots of each day; (**c**) Division of vehicle types; (**d**) Division of vehicle colors.

7. Challenges Regarding Vehicle Re-Identification

The vehicle re-id is among an essential and challenging task, and it is defined as, either any specific vehicle captured in one camera has already appeared over multiple camera network or not. With the increasing need for automated video analysis, the vehicle re-id receives increasing attention these days in the computer vision research community. Therefore, some key factor and their effects on performance are explained following.

- Insufficient data: For vehicle re-id systems each single image should match with gallery images, so it is very hard to get sufficient data for good model learning of each intra-class variability. However, it is also major challenge that dataset should reflect the real-world surveillance, currently, most of the datasets available are consists of non-overlapping views with a limited number of cameras; as a result, datasets have few viewpoints with unchanged regulation, and most of the publicly available datasets are consists of limited instances and classes that influence the performances.
- Inter-class similarity: This problem arises due to different automobile manufacturing companies have a similar visual appearance, as a result, two different make, model, and type of vehicles looks similar from rear or front side, as shown in Figure 27. [105,106].
- Intra-class variability: due to the unconstrained environment and viewpoint, the same vehicle looks different over different geographical locations of the surveillance camera network [107], as depicted in Figure 27.
- Pose and viewpoint variations: Due to the camera calibration, viewing angle and location on the roadside, captured vehicle image appearance varies, and the same vehicle looks different and different looks same. A learned model on the rear pose of a vehicle will probably fail to detect a vehicle's front, side pose. Furthermore, the effect of viewpoint change on vehicle is shown in Figure 28.

a) Intra-class variance   b) Inter-class similarity

Figure 27. Demonstration of two main challenges in vehicle re-id. (**a**) Intra-class variance; (**b**) Inter-class similarity.

Figure 28. Images of the same vehicle taken from different cameras to illustrate the appearance changes.

- Partial occlusions: If some part of an input vehicle is hidden by any object or vehicle in congestion as result, some key discriminative parts are not visible and the matching fails probably. Moreover, due to these features generated by an occluded vehicle image is corrupted [108].
- Illumination changes: Vehicle captured images illumination varies surveillance camera to surveillance camera and surveillance camera scenes and also illumination changes on the same surveillance camera due to different time slots like day and night. The same vehicle observed in different lighting conditions can have a color difference on the appearance because of the unconstrained environment [109]. Vehicles lights also have an effect on image illumination, so vehicle appearance changes at different a period of time and multiple camera network [110].
- Resolutions variation: Changes in resolution in pair of same vehicle occurs because of camera calibration, and another factor is various old surveillance cameras with different heights are fixed on the roadside that give a different-resolution,
- Deformation: Due to load or accident, vehicle shape, and body changes.

- Background clutter: This problem occurs in vehicle re-id when the vehicle's color and image background is the same.
- Changes in color response: The color attribute is one of the key parameters in vehicle re-id, but surveillance cameras color response changes because of camera settings features [110].
- Lighting effects: Specular reflection and shadows of the vehicle body generate the noise in vehicle image feature descriptor. If vehicle shadow is larger, there are more chance of inconsistency and noise in feature descriptor. As compared to the practical environment in a controlled environment, the lights and specular reflections can be controlled; but practically, we cannot control shadows, and it is one of the major problems in extracting information from the vehicle image
- Long-term re-id: If the same vehicle is captured after a long time or captured at different locations, then there is a high possibility that the vehicle looks different shape wise due to extra carry load/object.
- Cross dataset vehicle re-id: In vehicle re-id systems training and testing of model is performed on same dataset, but it is practically infeasible, due to significant difference between training and testing data and model may not generalize well.
- Insufficient temporal data: Due to the absence of unconstrained environmental information in datasets, it is impossible to exploit temporal data. However, temporal information can play an important role in the performance of vehicle re-id system.
- Vehicle re-id system scalability: Scalability means the system can adapt to varying factors while maintaining the performance, such as storing large gallery sizes that are constantly increasing and computational devices that efficiently analyze data.
- Real-time processing: Practical applications require real-time video processing, and the time constraint is the main challenge in vehicle re-id systems.
- Data labeling: This is a common difficulty in the computer vision field. Training a good model robust to all variations in a supervised way couldn't be done without a sufficient amount of annotated data. For a large camera network, manually collecting and annotating the amount of data from each surveillance camera is expensive.
- A small number of images per identity for training: Since one vehicle may appear very limited times in a camera network, it's difficult to collect much data of one single vehicle. Thus, usually data is insufficient to learn a good model of each specific vehicle's intra-class variability.
- A large number of candidates in gallery set: A camera network may cover a large public space, like a parking lot. Thus, there can be a huge amount of candidate for a given re-id query, and the number of candidates increases over time. The computation for matching with a large gallery set becomes expensive.
- Camera setting: Due to different camera settings and features, the same vehicle image captured by different cameras shows color dissimilarities. There may also be some geometric differences. For example, the shape of a vehicle may be observed with varying aspect ratios.
- Computation: All the proposed methods are based on deep learning. The computation for the training step with back propagation is more expensive than classical methods. In most cases, a powerful GPU is advisable for training, and more computation and memory resources are thus necessary. In applications with real-time constraints and without GPU, a very deep network may not be suitable for inference.

8. Evaluation Metrics

In the re-id task, the target object's images are mostly aligned and cropped. However, the vehicle re-id task is same as the instance retrieval. Given the input image, the candidates with a similar input image in the gallery set are required to be placed in the top positions within a ranking list. To measure the performance of vehicle re-id approaches, the cumulative matching characteristics (CMC), curve HIT@1 and HIT@5 are commonly used by researchers. CMC curve provides the probability that an input image identity appears

in a different-sized gallery set as shown in Figure 29. The cumulative number of correctly matched inputs is demonstrated based on the rank list in which inputs are re-identified. Moreover, HIT@1 is precision at rank-1 and HIT@5 is precision at rank-5. Rank is utilized to measure the matching score of test image to its own class, and higher value of rank indicates the improved performance of the system. Where the number of correctly re-identified input images in rank 1 is $q(i)$, the CMC value for rank i can be defined as:

$$CMC(i) = \sum_{r=1}^{i} q(r) \tag{4}$$

where r represents the rank index. CMC curve not only computes the rank-1 but also places the correctly matched images top ranks. Therefore, the CMC curve is a suitable option to describe the vehicle re-id performance of different approaches. Besides, CMC curves, if multiple ground truths for each query image in the gallery set are available, mean average precision (mAP) is used to measure the overall performance for vehicle re-id system. For the given query image, the average precision (AP) can be defined as:

$$AP = \frac{\sum_{k=1}^{n} P(k) \times G(k)}{Ngt} \tag{5}$$

where n is the number of test tracks and Ngt represents the number of ground truths. $P(k)$ shows the precision at cut-off k in the ranking lists. $G(k)$ equals 1 if the k-th match is true, otherwise 0. The mAP measures the overall performance of vehicle re-id system. Therefore, the mAP can be defined as:

$$mAP = \frac{\sum_{q=1}^{Q} AP(q)}{Q} \tag{6}$$

where Q denotes the number of queries.
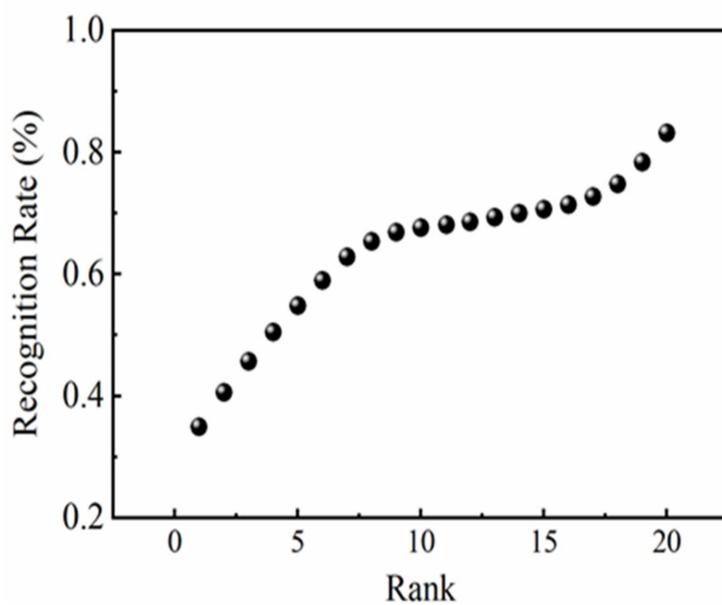

Figure 29. Cumulative matching characteristics (CMC) curve.

Another way by which vehicle re-id techniques performance can be evaluated is the confusion matrix. A confusion matrix consists of various columns and rows; it depends on the number of classes. It's diagonal represents the recognizing accuracy or true classification and off-diagonal express the misclassification.

9. Performance Comparison of Recently Proposed State-of-the-Art Approaches

The first phase in vehicle re-id is to decide whether the given vehicle image exists in the gallery set or not. In other terms, before considering for a similar match, the vehicle re-id system should have the capacity to decide whether the given vehicle probe image is a part of the gallery set or not. This approach is known novelty detection and it needs that vehicle re-id systems to have the ability to discard the miss-matched vehicle images. Usually in vehicle re-id systems, once the gallery set images are ranked in comparison with the given query image, the query image belongs to the gallery set if the similarity distance is higher than an operating threshold. We give a summary of the vehicle re-id mAP of some state-of-the-art methods including CMGN+Pre+Track [111], DF-CVTC [79], PROVID [42], and RAM [112] etc. on VeRi-776 dataset mentioned above. We have chosen VeRi-776 dataset for comparison because it is consisting of varying illumination, more viewpoints, and resolution. In short, this dataset fulfills most of the aspects of real-world camera surveillance data. The statistics about this dataset have been provided in Table 3.
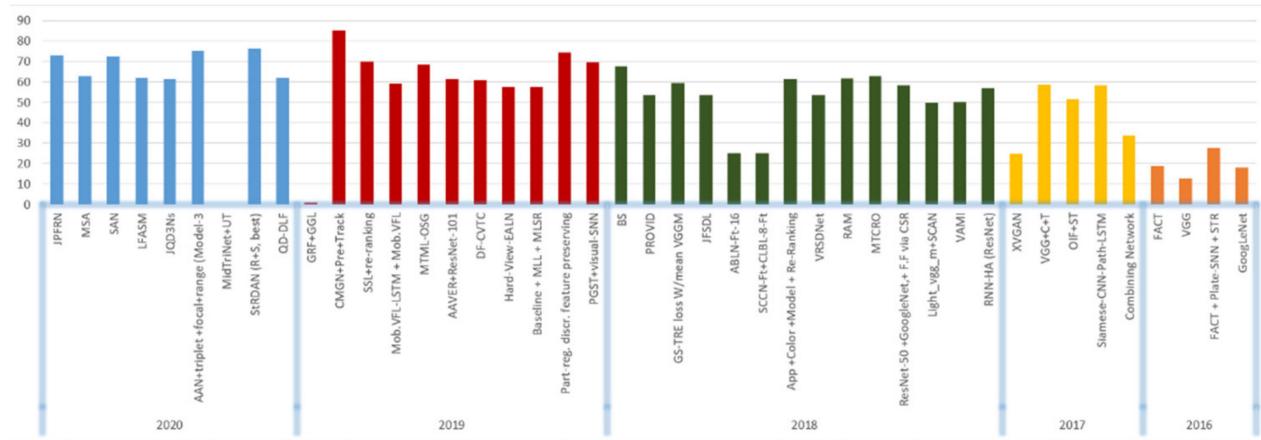
However, Table 4 provides recently proposed state-of-the-art approaches on VeRi-776 dataset. For comparison, we measure the performances of each method in mAP, HIT@1 and HIT@5. From Table 4, and Figure 30 we can observe that mAP of different models is increasing during the years 2016 to 2020. As on VeRi-776 dataset from the years 2016 to 2020, the performances of state-of-the-art methods have improved from 12.76% to 85.20%, with an increase of 72.44%. Moreover, Figure 31 shows the CMC of different state-of-the-art approaches on VehicleID dataset with different test size.

Table 4. Performance analysis of some proposed approaches in state-of-art on VeRi-776.

Reference	Venue	Approach	mAP	HIT-1%	HIT5%
		Year 2020			
L. Xiangwei et al. [113]	Mobile Networks and Applications	JPFRN	72.86	93.14	97.85
Z. Aihua et al. [114]	Neural Computing and Applications	MSA	62.89	92.07	96.19
Q. Jingjing et al. [115]	Measurement Science and Technology	SAN	72.5	93.3	97.1
W. Honglie et al. [116]	Applied Sciences	LFASM	61.92	90.11	92.91
Z. Jianqing et al. [117]	IEEE Internet of Things	JQD^{3}Ns	61.30	89.69	95.17
Z. Hui et al. [118]	IEEE ITNEC	AAN+triplet +focal+range (Model-3	75.14	5.17	97.80
O. Daniel et al. [119]	IEEE Access	MidTriNet+UT	......	89.15	93.74
L. Sangroket al. [120]	CVPRW	StRDAN (R+S, best)	76.1	......	......
J. Zhu et al. [121]	IEEE TITS	QD-DLF	61.83	88.50	94.46
L. Xiaobin et.al. [122]	IEEE Trans. on Image Processing	GRF+GGL	00.61	0.89	0.95
		Year 2019			
A. A-Acevedo et al. [111]	IEEECVPR	CMGN+Pre+Track	85.20	96.60	......
F. Wu et al. [123]	Image Communication	SSL+re-ranking	69.90	89.69	95.41
S. Ahmed et al. [124]	IEEE ICIP	Mob.VFL-LSTM + Mob.VFL	59.18	88.08	94.63
G. Rajamanoharan et al. [125]	IEEE CVPR	MTML-OSG	68.3	92.0	94.2
P. Khorram et al. [88]	ArXiv	AAVER+ResNet-101	61.18	88.97	94.70
A. Zheng et al. [79]	ArXiv	DF-CVTC	61.06	91.36	95.77
Y. Lou et al. [74]	IEEE TIP	Hard-View-EALN	57.44	84.39	94.05
J. Hou et al. [126]	Neurocomputing	Baseline + MLL + MLSR	57.52	87.19	94.16
B. He et al. [127]	IEEE CVPR	Part-reg. discr. feature preserving	74.3	94.3	98.7
X. Zhong et al. [128]	ICMM	PGST+visual-SNN	69.47	89.36	94.40
R. Kumar et al. [107]	IJCNN	BS	67.55	90.23	96.42

Table 4. *Cont.*

| Reference	Venue	Approach	mAP	HIT-1%	HIT5%
		Year 2018			
X. Liu et al. [42]	IEEE Trans. on Multimedia	PROVID	53.42	81.56	95.11
Y. Bai et al. [8]	IEEE Trans. on Multimedia	GS-TRE loss W/mean VGGM	59.47	96.24	98.97
J. Zhu et al. [129]	IEEE Access	JFSDL	53.53	82.90	91.60
Y. Zhou et al. [70]	IEEE WACV	ABLN-Ft-16	24.92	60.49	77.33
Y. Zhou et al. [69]	IEEE TIP	SCCN-Ft+CLBL-8-Ft	25.12	60.83	78.55
N. Jiang et al. [98]	IEEE ICIP	App +Color +Model + Re-Ranking	61.11	89.27	94.76
J. Zhu et al. [130]	MM Tools and Applications	VRSDNet	53.45	83.49	92.55
X. Liu et al. [112]	IEEE IME	RAM	61.5	88.6	94.0
D. Xu et al. [110]	ICIMCS	MTCRO	62.61	87.96	94.63
D. Sun et al. [131]	Springer ICBICS	ResNet-50 +GoogleNet,+ F.F via CSR	58.21	90.52	93.38
S. Teng et al. [87]	Springer PCM	Light_vgg_m+SCAN	49.87	82.24	90.76
Y. Zhou et al. [83]	CVPR	VAMI	50.13	77.03	90.82
Xiu-Shen et al. [86]	ACCV	RNN-HA (ResNet)	56.80	74.79	87.31
		Year 2017			
Y. Zhou et al. [78]	BMVC	XVGAN	24.65	60.20	77.03
Y. zhang et al. [46]	IEEE ICME	VGG+C+T	58.78	86.41	92.91
Z. Wang et al. [67]	ICCV	OIF+ST	51.4	92.35	......
Y. Shen et al. [97]	ICCV	Siamese-CNN-Path-LSTM	58.27	83.49	90.04
Y. Tang et al. [132]	IEEE ICIP	Combining Network	33.78	60.19	77.40
		Year 2016			
X. Liu et al. [43]	IEEE ICME	FACT	18.75	52.21	72.88
H. Liu et al. [44]	CVPR	VGG	12.76	44.10	62.63
X. Liu et al. [96]	ECCV	FACT + Plate-SNN + STR	27.77	61.44	78.78
L. Yang et al. [101]	CVPR	GoogLeNet	17.89	52.32	72.17

Figure 30. Demonstrates the performance comparison of different state of the art approaches.
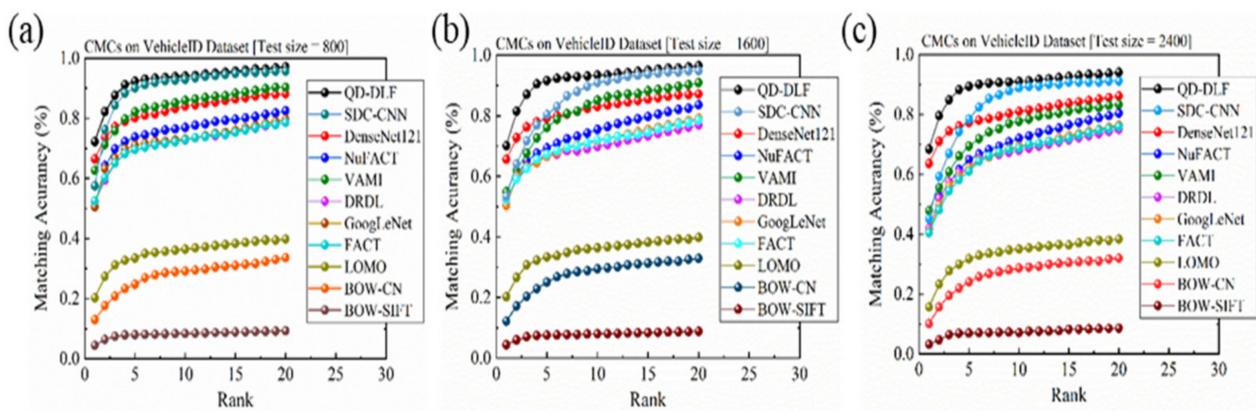
Figure 31. Demonstrates the performance comparison of different state-of-the-art approaches on VehicleID dataset. (**a**) Test size = 800; (**b**) Test size = 1600; (**c**) Test size = 2400.

10. Conclusions & Way Forward

Vehicle re-id is one of the most critical and challenging area in the ITS. Despite high significance, it is not well explored compared to a similar problem that is person re-id. In this review paper, the authors present recent advancements being done for vehicle re-id. Moreover, to draw a detailed picture of study, the authors discuss different vehicle re-id technologies, especially vision-based, including appearance, license plate, spatio-temporal, etc., along with the quantitative and qualitative comparison of different vision-based methods on VeRi-776 and VehicleID datasets. In addition, this review provides comprehensive synopses of publicly available benchmark datasets utilized for performance evaluation with a brief description of re-id evaluation techniques. This paper also presents applications as well as the main challenges of camera-based vehicle re-id such as complex and unconstrained environment, dirt, snow, occluded image, blurry image, and sunshine, etc., along with varied road topology that affects the performance.

There are many aspects of vehicle re-id that can be improved. In the future, a reader can explore possibilities to enhance the overall performance of vehicle re-id. Moreover, there is significant potential to extend the approach with some of the following concepts:

CNN works on edges, shapes, and original vehicle features, but the relationship between these features is not considered; hence, the model performance is often unsatisfactory when the vehicle image is rotated or captured with a different rotation. However, a recent capsule network [133] is introduced, which showed improved performance in handling different poses, orientations, and occluded objects.

Secondly, attention-based deep neural network models have gained encouraging results on various challenging tasks, including machine translation [134], caption generation [135], and object recognition [136]. However, attention-based neural network models are still not well investigated for vehicle re-id.

Lastly, due to the development of large-scale real-world data sets, the vehicle re-id system's performance is significantly increased. However, existing datasets offer a specific range of vehicle images with correlated data that causes over-fitting due to over-tuned parameters on specific data. Therefore, the system cannot efficiently generalize other data. A reader can develop large scale real-world surveillance vehicle datasets in an unconstrained environment with multiple views to enhance the training of the state-of-the-art approaches for performance improvement.

Concisely, Vehicle re-id is a demanding and challenging area with massive opportunities for improvement and research. This review paper attempts to provide an overview of the vehicle re-id problem, its challenges, and applications, and, simultaneously, present a way forward. We hope this paper will be valuable for anyone who wants to work in this area.

References

1. The Ministry of Transport of the People's Republic of China. *Annual Report of City Transportation Developmentin China*; The Ministry of Transport of the People's Republic of China: Beijing, China, 2010.
2. Chen, Z.; Fan, W.; Xiong, Z.; Zhang, P.; Luo, L. Visual Data Security and Management for Smart Cities. *Front. Comput. Sci. China* **2010**, *4*, 386–393. [CrossRef]
3. Cheng, K.; Khokhar, M.S.; Ayoub, M.; Jamali, Z. Nonlinear Dimensionality Reduction in Robot Vision for Industrial Monitoring Process via Deep Three Dimensional Spearman Correlation Analysis (D3D-SCA). *Multimed. Tools Appl.* **2020**, *80*, 5997–6017. [CrossRef]
4. Khokhar, M.S.; Cheng, K.; Ayoub, M.; Eric, L.K. Multi-Dimension Projection for Non-Linear Data Via Spearman Correlation Analysis (MD-SCA). In Proceedings of the 2019 8th International Conference on Information and Communication Technologies (ICICT), Karachi, Pakistan, 16–17 November 2019; pp. 14–18.
5. Qian, Q.; Jin, R.; Zhu, S.; Lin, Y. Fine-Grained Visual Categorization via Multi-Stage Metric Learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; IEEE: Piscataway, NJ, USA; pp. 3716–3724.
6. Cui, Y.; Zhou, F.; Lin, Y.; Belongie, S. Fine-Grained Categorization and Dataset Bootstrapping Using Deep Metric Learning with Humans in the Loop. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; IEEE: Piscataway, NJ, USA; pp. 1153–1162.
7. Zhao, Y.; Shen, C.; Wang, H.; Chen, S. Structural Analysis of Attributes for Vehicle Re-Identification and Retrieval. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 723–734. [CrossRef]
8. 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]
9. Deng, J.; Cai, J.; Aftab, M.U.; Khokhar, M.S.; Kumar, R. Visual Features with Spatio-Temporal-Based Fusion Model for Cross-Dataset Vehicle Re-Identification. *Electronics* **2020**, *9*, 1083. [CrossRef]
10. Wu, Y.; Lin, Y.; Dong, X.; Yan, Y.; Bian, W.; Yang, Y. Progressive Learning for Person Re-Identification with One Example. *IEEE Trans. Image Process.* **2019**, *28*, 2872–2881. [CrossRef] [PubMed]
11. Rasouli, A.; Kotseruba, I.; Tsotsos, J.K. Understanding Pedestrian Behavior in Complex Traffic Scenes. *IEEE Trans. Intell. Veh.* **2018**, *3*, 61–70. [CrossRef]
12. Chen, Z.; Liao, W.; Xu, B.; Liu, H.; Li, Q.; Li, H.; Xiao, C.; Zhang, H.; Li, Y.; Bao, W.; et al. Object Tracking over a Multiple-Camera Network. In Proceedings of the IEEE International Conference on Multimedia Big Data, Beijing, China, 20–22 April 2015; pp. 276–279.
13. Cai, J.; Deng, J.; Khokhar, M.S.; Umar Aftab, M. Vehicle Classification Based on Deep Convolutional Neural Networks Model for Traffic Surveillance Systems. In Proceedings of the 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 14–16 December 2018; pp. 224–227.
14. Arandjelovic, R.; Zisserman, A. Three Things Everyone Should Know to Improve Object Retrieval. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2911–2918.
15. Khokhar, M.S.; Cheng, K.; Ayoub, M.; Rub, N.E. Data Driven Processing Via Two-Dimensional Spearman Correlation Analysis (2D-SCA). In Proceedings of the 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), Karachi, Pakistan, 14–15 December 2019; pp. 1–7.
16. Yan, L.; Wang, Y.; Song, T.; Yin, Z. An Incremental Intelligent Object Recognition System Based on Deep Learning. In Proceedings of the 2017 Chinese Automation Congress (CAC), Jinan, China, 20–22 October 2017; pp. 7135–7138.

17. Saghafi, M.A.; Hussain, A.; Saad, M.H.M.; Tahir, N.M.; Zaman, H.B.; Hannan, M.A. Appearance-Based Methods in Re-Identification: A Brief Review. In Proceedings of the 2012 IEEE 8th International Colloquium on Signal Processing and its Applications, Malacca, Malaysia, 23–25 March 2012; pp. 404–408.

18. Doretto, G.; Sebastian, T.; Tu, P.; Rittscher, J. Appearance-Based Person Reidentification in Camera Networks: Problem Overview and Current Approaches. *J. Ambient. Intell. Humaniz. Comput.* **2011**, *2*, 127–151. [CrossRef]

19. Yi, D.; Lei, Z.; Li, S.Z. Deep Metric Learning for Practical Person Re-Identification. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 37–39. [CrossRef]

20. Kwong, K.; Kavaler, R.; Rajagopal, R.; Varaiya, P. A Practical Scheme for Arterial Travel Time Estimation Based on Vehicle Re-Identification Using Wireless Sensors. *Transp. Res. Part C Emerg. Technol.* **2009**, *17*, 586–606. [CrossRef]

21. Wang, R.; Zhang, L.; Sun, R.; Gong, J.; Cui, L. Easitia: A Pervasive Traffic Information Acquisition System Based on Wireless Sensor Networks. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 615–621. [CrossRef]

22. Sart, D.; Mueen, A.; Najjar, W.; Keogh, E.; Niennattrakul, V. Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs. In Proceedings of the 2010 IEEE International Conference on Data Mining, Sydney, NSW, Australia, 13–17 December 2010; pp. 1001–1006.

23. Tarango, J.; Keogh, E.; Brisk, P. Instruction Set Extensions for Dynamic Time Warping. In Proceedings of the 2013 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), Montreal, QC, Canada, 29 September–4 October 2013; pp. 1–10.

24. Charbonnier, S.; Pitton, A.-C.; Vassilev, A. Vehicle Re-Identification with a Single Magnetic Sensor. In Proceedings of the 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Graz, Austria, 13–16 May 2012; pp. 380–385.

25. Sanchez, R.O.; Flores, C.; Horowitz, R.; Rajagopal, R.; Varaiya, P. Vehicle Re-Identification Using Wireless Magnetic Sensors: Algorithm Revision, Modifications and Performance Analysis. In Proceedings of the 2011 IEEE International Conference on Vehicular Electronics and Safety, Beijing, China, 10–12 July 2011; pp. 226–231.

26. Jeng, S.C.; Chu, L. Vehicle Reidentification with the Inductive Loop Signature Technology. *J. East. Asia Soc. Transp. Stud.* **2013**, *10*, 1896–1915. [CrossRef]

27. Sun, C.; Ritchie, S.G.; Tsai, K.; Jayakrishnan, R. Use of Vehicle Signature Analysis and Lexicographic Optimization for Vehicle Reidentification on Freeways. *Transp. Res. Part C Emerg. Technol.* **1999**, *7*, 167–185. [CrossRef]

28. Jeng, S.-T.; Tok, Y.C.A.; Ritchie, S.G. Freeway Corridor Performance Measurement Based on Vehicle Reidentification. *IEEE Trans. Intell. Transp. Syst.* **2010**, *11*, 639–646. [CrossRef]

29. Kwon, T.M. Blind Deconvolution of Vehicle Inductance Signatures for Travel-Time Estimation. In Proceedings of the 85th Annual Meeting of the Transportation Research Board, Washington, DC, USA, January 22–25 2006; Elsevier: Amsterdam, The Netherlands; pp. 1–12.

30. Blokpoel, R.J. Vehicle Reidentification Using Inductive Loops in Urban Areas. In Proceedings of the 16th ITS World Congress and Exhibition on Intelligent Transport Systems and Services, Stockholm, Sweden, 21–25 September 2009; pp. 1–7.

31. Kim, J.-H.; Oh, J.-H. A Land Vehicle Tracking Algorithm Using Stand-Alone GPS. *Control Eng. Pract.* **2000**, *8*, 1189–1196. [CrossRef]

32. Taylor, S.Y.; Green, J.; Richardson, A.J. *Applying Vehicle Tracking and Palmtop Technology to Urban Freight Surveys*; Transport Data Centre, Dept. of Transport NSW: New South Wales, Australia, 1998.

33. Sun, Z.; Ban, X. (Jeff) Vehicle Classification Using GPS Data. *Transp. Res. Part C Emerg. Technol.* **2013**, *37*, 102–117. [CrossRef]

34. Wang, X. Intelligent Multi-Camera Video Surveillance: A Review. *Pattern Recognit. Lett.* **2013**, *34*, 3–19. [CrossRef]

35. Sochor, J.; Spanhel, J.; Herout, A. Boxcars: Improving Fine-Grained Recognition of Vehicles Using 3-D Bounding Boxes in Traffic Surveillance. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 97–108. [CrossRef]

36. Zheng, L.; Shen, L.; Tian, L.; Wang, S.; Wang, J.; Tian, Q. Scalable Person Re-Identification: A Benchmark. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; IEEE: Piscataway, NJ, USA; pp. 1116–1124.

37. Liao, S.; Hu, Y.; Zhu, X.; Li, S.Z. Person Re-Identification by Local Maximal Occurrence Representation and Metric Learning. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; IEEE: Piscataway, NJ, USA; pp. 2197–2206.

38. 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, June 7–12 2015; IEEE: Piscataway, NJ, USA; pp. 1–9.

39. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556. [CrossRef]

40. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

41. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

42. 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]

43. Liu, X.; Liu, W.; Ma, H.; Fu, H. Large-Scale Vehicle Re-Identification in Urban Surveillance Videos. In Proceedings of the 2016 IEEE International Conference on Multimedia and Expo (ICME), Seattle, WA, USA, 11–15 July 2016; pp. 1–6.

44. 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; pp. 2167–2175.

45. Li, Y.; Li, Y.; Yan, H.; Liu, J. Deep Joint Discriminative Learning for Vehicle Re-Identification and Retrieval. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 395–399.

46. 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.

47. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up Robust Features. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417.

48. Lowe, G.D. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]

49. Guo, G.; Wang, H.; Bell, D.; Bi, Y.; Greer, K. KNN Model-Based Approach in Classification. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 986–996.

50. Gunn, S.R. Support Vector Machines for Classification and Regression. 1998. Available online: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.7171&rep=rep1&type=pdf (accessed on 10 May 1998).

51. Friedman, N.; Geiger, D.; Goldszmidt, M. Bayesian Network Classifiers. *Mach. Learn.* **1997**, *29*, 131–163. [CrossRef]

52. Quinlan, J.R. Induction of Decision Trees. *Mach. Learn.* **1986**, *1*, 81–106. [CrossRef]

53. 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 (CVPRW), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1568–1574.

54. Chen, H.C.; Hsieh, J.-W.; Huang, S.-P. Real-Time Vehicle Re-Identification System Using Symmelets and HOMs. In Proceedings of the 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand, 27–30 November 2018; pp. 1–6.

55. Cormier, M.; Sommer, L.W.; Teutsch, M. Low Resolution Vehicle Re-Identification Based on Appearance Features for Wide Area Motion Imagery. In Proceedings of the 2016 IEEE Winter Applications of Computer Vision Workshops (WACVW), Lake Placid, NY, USA, 10 March 2016; pp. 1–7.

56. Yang, L.; Jin, R. *Distance Metric Learning: A Comprehensive Survey*; Michigan State University: Michigan, MI, USA, 2006.

57. Sun, Y.; Wang, X.; Tang, X. Deep Learning Face Representation from Predicting 10,000 Classes. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1891–1898.

58. Sun, Y.; Wang, X.; Tang, X. Deep Learning Face Representation by Joint Identification-Verification. In Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14), Montreal, QC, Canada, 8–13 December 2014; pp. 1988–1996.

59. Zhang, L.; Xiang, T.; Gong, S. Learning a Discriminative Null Space for Person Re-Identification. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1239–1248.

60. Shih, K.J.; Mallya, A.; Singh, S.; Hoiem, D. Part Localization Using Multi-Proposal Consensus for Fine-Grained Categorization. *arXiv* **2015**, arXiv:1507.06332.

61. Liu, X.; Xia, T.; Wang, J.; Yang, Y.; Zhou, F.; Lin, Y. Fully Convolutional Attention Networks for Fine-Grained Recognition. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), San Francisco, CA, USA, 22 March 2016; pp. 4190–4196.

62. Lin, T.-Y.; RoyChowdhury, A.; Maji, S. Bilinear CNN Models for Fine-Grained Visual Recognition. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1449–1457.

63. Boonsim, N.; Prakoonwit, S. Car Make and Model Recognition under Limited Lighting Conditions at Night. *Pattern Anal. Appl.* **2017**, *20*, 1195–1207. [CrossRef]

64. Zhao, H.; Tian, M.; Sun, S.; Shao, J.; Yan, J.; Yi, S.; Wang, X.; Tang, X. Spindle Net: Person Re-Identification with Human Body Region Guided Feature Decomposition and Fusion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, NJ, USA; pp. 907–915.

65. Wu, Z.; Li, Y.; Radke, R.J. Viewpoint Invariant Human Re-Identification in Camera Networks Using Pose Priors and Subject-Discriminative Features. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1095–1108. [CrossRef]

66. Zheng, L.; Huang, Y.; Lu, H.; Yang, Y. Pose-Invariant Embedding for Deep Person Re-Identification. *IEEE Trans. Image Process.* **2019**, *28*, 4500–4509. [CrossRef] [PubMed]

67. 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 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; IEEE: Piscataway, NJ, USA; pp. 379–387.

68. Prokaj, J.; Medioni, G. 3-D Model Based Vehicle Recognition. In Proceedings of the 2009 Workshop on Applications of Computer Vision (WACV), Snowbird, UT, USA, 7–8 December 2009; pp. 1–7.

69. 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]

70. 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.

71. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the International Conference on Neural Information Processing Systems, Montreal, Canada, 8–13 December 2014; MIT Press: Cambridge, MA, USA; pp. 2672–2680.

72. 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 Reidentification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; IEEE: Piscataway, NJ, USA, 2018.

73. Radford, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv* **2015**, arXiv:1511.06434.

74. Lou, Y.; Bai, Y.; Liu, J.; Wang, S.; Duan, L.-Y. Embedding Adversarial Learning for Vehicle Re-Identification. *IEEE Trans. Image Process.* **2019**, *28*, 3794–3807. [CrossRef] [PubMed]

75. Ma, L.; Sun, Q.; Georgoulis, S.; Van Gool, L.; Schiele, B.; Fritz, M. Disentangled Person Image Generation. In Proceedings of the In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 99–108.

76. Zheng, Z.; Zheng, L.; Yang, Y. Unlabeled Samples Generated by Gan Improve the Person Re-Identification Baseline in Vitro. *arXiv Preprint* **2017**, arXiv:1701.077173.

77. Wei, L.; Zhang, S.; Gao, W.; Tian, Q. Person Transfer Gan to Bridge Domain Gap for Person Re-Identification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; IEEE: Piscataway, NJ, USA, 2018.

78. Zhou, Y.; Shao, L. Cross-View GAN Based Vehicle Generation for Re-Identification. In Proceedings of the British Machine Vision Conference, London, UK, 4–7 September 2017; BMVA Press: London, UK, 2017; pp. 1–12.

79. Zheng, A.; Lin, X.; Li, C.; He, R.; Tang, J. Attributes Guided Feature Learning for Vehicle Re-Identification. *arXiv* **2019**, arXiv:1905.08997.

80. Wang, F.; Jiang, M.; Qian, C.; Yang, S.; Li, C.; Zhang, H.; Wang, X.; Tang, X. Residual Attention Network for Image Classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, NJ, USA; pp. 6450–6458.

81. Fu, J.; Zheng, H.; Mei, T. Look Closer to See Better: Recurrent Attention Convolutional Neural Network for Fine-Grained Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, NJ, USA; pp. 4476–4484.

82. Sharma, S.; Kiros, R.; Salakhutdinov, R. Action Recognition Using Visual Attention. *arXiv* **2015**, arXiv:1511.04119.

83. Zhouy, Y.; Shao, L. Viewpoint-Aware Attentive Multi-View Inference for Vehicle Re-Identification. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; IEEE: Piscataway, NJ, USA; pp. 6489–6498.

84. Zhao, L.; Li, X.; Zhuang, Y.; Wang, J. Deeply-Learned Part-Aligned Representations for Person Re-Identification. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; IEEE: Piscataway, NJ, USA; pp. 3239–3248.

85. Guo, H.; Zhu, K.; Tang, M.; Wang, J. Two-Level Attention Network with Multi-Grain Ranking Loss for Vehicle Re-Identification. *IEEE Trans. Image Process.* **2019**, *28*, 4328–4338. [CrossRef]

86. Wei, X.-S.; Zhang, C.-L.; Liu, L.; Shen, C.; Wu, J. Coarse-to-Fine: A RNN-Based Hierarchical Attention Model for Vehicle Re-Identification. In Proceedings of the 14th Asian Conference on Computer Vision, Perth, Australia, 2–6 December 2018; pp. 574–591.

87. Teng, S.; Liu, X.; Zhang, S.; Huang, Q. SCAN: Spatial and Channel Attention Network for Vehicle Re-Identification. In *Pacific Rim Conference on Multimedia*; Springer International Publishing: Cham, Switzerland, 2018; pp. 350–361.

88. Khorramshahi, P.; Kumar, A.; Peri, N.; Rambhatla, S.S.; Chen, J.-C.; Chellappa, R. A Dual-Path Model with Adaptive Attention for Vehicle Re-Identification. *arXiv* **2019**, arXiv:1905.03397.

89. Zhang, X.; Zhang, R.; Cao, J.; Gong, D.; You, M.; Shen, C. Part-Guided Attention Learning for Vehicle Re-Identification. *arXiv* **2019**, arXiv:1909.06023.

90. Qadri, M.T.; Asif, M. Automatic Number Plate Recognition System for Vehicle Identification Using Optical Character Recognition. In Proceedings of the 2009 International Conference on Education Technology and Computer, Singapore, 17–20 April 2009; pp. 335–338.

91. Li, H.; Shen, C. Reading Car License Plates Using Deep Convolutional Neural Networks and LSTMs. *arXiv* **2016**, arXiv:1601.05610.

92. Shi, B.; Bai, X.; Yao, C. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2298–2304. [CrossRef]

93. Ellis, T.; Makris, D.; Black, J.; Engineers, E. Learning a Multi-Camera Topology. In Proceedings of the Joint IEEE Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS), Nice, France, 11–12 October 2003; pp. 165–171.

94. Wu, C.W.; Liu, C.T.; Chiang, C.E.; Tu, W.C.; Chien, S.Y. Vehicle Re-Identification with the Space-Time Prior. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; IEEE: Piscataway, NJ, USA; pp. 121–128.

95. Loy, C.C.; Xiang, T.; Gong, S. Multi-Camera Activity Correlation Analysis. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 1988–1995.

96. Liu, X.; Liu, W.; Mei, T.; Ma, H. A Deep Learning-Based Approach to Progressive Vehicle Re-Identification for Urban Surveillance. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2016; pp. 869–884.

97. 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; pp. 1918–1927.

98. 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.

99. Jamali, Z.; Deng, J.; Cai, J.; Aftab, M.U.; Hussain, K. Minimizing Vehicle Re-Identification Dataset Bias Using Effective Data Augmentation Method. In Proceedings of the 2019 15th International Conference on Semantics, Knowledge and Grids (SKG), Guangzhou, China, 17–18 September 2019; pp. 127–130.

100. Guo, H.; Zhao, C.; Liu, Z.; Wang, J.; Lu, H. Learning Coarse-to-Fine Structured Feature Embedding for Vehicle Re-Identification. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence AAAI18, New Orleans, LA, USA, 2–7 February 2018; pp. 6853–6860.

101. 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; IEEE: Piscataway, NJ, USA; pp. 3973–3981.

102. Kanacı, A.; Zhu, X.; Gong, S. Vehicle re-identification in context. In *German Conference on Pattern Recognition*; Springer: Cham, Switzerland, 2018; pp. 377–390. [CrossRef]

103. Li, X.; Yuan, M.; Jiang, Q.; Li, G. VRID-1: A Basic Vehicle Re-Identification Dataset for Similar Vehicles. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 1–8.

104. Lou, Y.; Bai, Y.; Liu, J.; Wang, S.; Duan, L.-Y. VERI-Wild: A Large Dataset and a New Method for Vehicle Re-Identification in the Wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 3230–3238.

105. De Oliveira, I.O.; Fonseca, K.V.O.; Minetto, R. A Two-Stream Siamese Neural Network for Vehicle Re-Identification by Using Non-Overlapping Cameras. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 669–673.

106. Cai, J.; Deng, J.; Aftab, M.; Khokhar, M.; Kumar, R. Efficient and Deep Vehicle Re-Identification Using Multi-Level Feature Extraction. *Appl. Sci.* **2019**, *9*, 1291. [CrossRef]

107. Kumar, R.; Weill, E.; Aghdasi, F.; Sriram, P. Vehicle Re-Identification: An Efficient Baseline Using Triplet Embedding. *arXiv* **2019**, arXiv:1901.01015. [CrossRef]

108. 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.

109. Xu, D.; Lang, C.; Feng, S.; Wang, T. A Framework with a Multi-Task CNN Model Joint with a Re-Ranking Method for Vehicle Re-Identification. In Proceedings of the 10th International Conference on Internet Multimedia Computing and Service-ICIMCS'18, Nanjing, China, 17–19 August 2018; ACM Press: New York, NY, USA, 2018; pp. 1–7.

110. 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]

111. Ayala-acevedo, A.; Devgun, A.; Zahir, S.; Askary, S. Vehicle Re-Identification: Pushing the Limits of Re-Identification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; IEEE: Piscataway, NJ, USA; 2019; pp. 291–296.

112. 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.

113. Lin, X.; Zeng, H.; Hou, J.; Cao, J.; Zhu, J.; Chen, J. Joint Pyramid Feature Representation Network for Vehicle Re-Identification. *Mob. Netw. Appl.* **2020**, *25*, 1781–1792. [CrossRef]

114. Zheng, A.; Lin, X.; Dong, J.; Wang, W.; Tang, J.; Luo, B. Multi-Scale Attention Vehicle Re-Identification. *Neural Comput. Appl.* **2020**, *32*, 17489–17503. [CrossRef]

115. Qian, J.; Jiang, W.; Luo, H.; Yu, H. Stripe-Based and Attribute-Aware Network: A Two-Branch Deep Model for Vehicle Re-Identification. *Meas. Sci. Technol.* **2020**, *31*, 95401. [CrossRef]

116. Wang, H.; Sun, S.; Zhou, L.; Guo, L.; Min, X.; Li, C. Local Feature-Aware Siamese Matching Model for Vehicle Re-Identification. *Appl. Sci.* **2020**, *10*, 2474. [CrossRef]

117. Zhu, J.; Huang, J.; Zeng, H.; Ye, X.; Li, B.; Lei, Z.; Zheng, L. Object Reidentification via Joint Quadruple Decorrelation Directional Deep Networks in Smart Transportation. *IEEE Internet Things J.* **2020**, *7*, 2944–2954. [CrossRef]

118. Zhou, H.; Li, C.; Zhang, L.; Song, W. Attention-Aware Network and Multi-Loss Joint Training Method for Vehicle Re-Identification. In Proceedings of the 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, 12–14 June 2020; pp. 1330–1334.

119. Organisciak, D.; Sakkos, D.; Ho, E.S.L.; Aslam, N.; Shum, H.P.H. Unifying Person and Vehicle Re-Identification. *IEEE Access* **2020**, *8*, 115673–115684. [CrossRef]

120. Lee, S.; Park, E.; Yi, H.; Lee, S.H. StRDAN: Synthetic-to-Real Domain Adaptation Network for Vehicle Re-Identification. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Virtual Conference, 14–19 June 2020; IEEE: Piscataway, NJ, USA; pp. 2590–2597.

121. Zhu, J.; Zeng, H.; Huang, J.; Liao, S.; Lei, Z.; Cai, C.; Zheng, L. Vehicle Re-Identification Using Quadruple Directional Deep Learning Features. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 410–420. [CrossRef]

122. Liu, X.; Zhang, S.; Wang, X.; Hong, R.; Tian, Q. Group-Group Loss-Based Global-Regional Feature Learning for Vehicle Re-Identification. *IEEE Trans. Image Process.* **2020**, *29*, 2638–2652. [CrossRef] [PubMed]

123. Wu, F.; Yan, S.; Smith, J.S.; Zhang, B. Vehicle Re-Identification in Still Images: Application of Semi-Supervised Learning and Re-Ranking. *Signal Process. Image Commun.* **2019**, 261–271. [CrossRef]

124. Alfasly, S.A.S.; Hu, Y.; Liang, T.; Jin, X.; Zhao, Q.; Liu, B. Variational Representation Learning for Vehicle Re-Identification. *arXiv* **2019**, arXiv:1905.02343v1. [CrossRef]

125. Rajamanoharan, G.; Kanacı, A.; Li, M.; Gong, S. Multi-Task Mutual Learning for Vehicle Re-Identification. In Proceedings of the The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; IEEE: Piscataway, NJ, USA; pp. 62–70.

126. Hou, J.; Zeng, H.; Cai, L.; Zhu, J.; Chen, J.; Ma, K.-K. Multi-Label Learning with Multi-Label Smoothing Regularization for Vehicle Re-Identification. *Neurocomputing* **2019**, *345*, 15–22. [CrossRef]

127. He, B.; Li, J.; Zhao, Y.; Tian, Y. Part-Regularized near-Duplicate Vehicle Re-Identification. In Proceedings of the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; IEEE: Piscataway, NJ, USA; pp. 3997–4005.

128. Zhong, X.; Feng, M.; Huang, W.; Wang, Z.; Satoh, S. Poses Guide Spatiotemporal Model for Vehicle Re-Identification. In *International Conference on Multimedia Modeling*; Springer International Publishing: Cham, Switzerland, 2019; pp. 427–440.

129. 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]

130. Zhu, J.; Du, Y.; Hu, Y.; Zheng, L.; Cai, C. VRSDNet: Vehicle Re-Identification with a Shortly and Densely Connected Convolutional Neural Network. *Multimed. Tools Appl.* **2018**, *78*, 29043–29057. [CrossRef]

131. 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 *International Conference on Brain Inspired Cognitive Systems*; Springer International Publishing: Cham, Switzerland, 2018; pp. 54–63.

132. Tang, Y.; Wu, D.; Jin, Z.; Zou, W.; Li, X. Multi-Modal 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; pp. 2254–2258.

133. Sabour, S.; Frosst, N.; Hinton, G.E. Dynamic Routing between Capsules. In Proceedings of the Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 26 October 2017; pp. 3859–3869.

134. Bahdanau, D.; Cho, K.; Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv* **2014**, arXiv:1409.0473.

135. Xu, K.; Ba, J.; Kiros, R.; Cho, K.; Courville, A.; Salakhutdinov, R.; Zemel, R.; Bengio, Y. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 10 February 2015; pp. 2048–2057.

136. Ba, J.; Mnih, V.; Kavukcuoglu, K. Multiple Object Recognition with Visual Attention. *arXiv* **2014**, arXiv:1412.7755.