A Survey on Vehicle Re-Identifcation

Harsh Bansal
The LNM institute of
Information Technology,
Jaipur, India
Email: 21ucc044@lnmiit.ac.in

Prianshu Shaw
The LNM institute of
Information Technology,
Jaipur, India
Email: 21ucs161@lnmiit.ac.in

Kanchi Sharma
JK Lakshmipat
University,
Jaipur, India
Email: kanchisharma@jklu.edu.in

With the intensive strides in the edge-cutting technologies in Computer Vision and Artificial Intelligence, Intelligent Transportation system helps the smart cities to automate the Vehicle-re-identification process using Computer Vision, Deep Learning and various other automation processes. This survey paper explores the methodologies used in vehicle re-identification (re-ID) to address the challenges of matching and identifying vehicles across multiple cameras. It discusses the limitations of traditional systems and highlights the advancements enabled by deep learning models, fusion techniques, attention mechanisms, and regularization techniques. Various datasets used in vehicle re-ID research are examined, along with evaluation strategies such as mean average precision (mAP) and Cumulative Matching Characteristic (CMC) curves. The paper presents the results and discussions of different approaches, showcasing their strengths and performance on benchmark datasets. It concludes by suggesting future research directions, including the integration of spatiotemporal information, multi-view re-ID, and the application of transfer learning to real traffic scenes. This survey serves as a comprehensive resource for researchers in the field of vehicle re-identification. It also analyses the different methology researched and deployed for vehicle re-ID, and compared their achievements and shortcomings in brief.

1 Introduction

In recent years, significant strides have been made in the fields of computer vision and Internet of Things (IoT), leading to diverse research and technological advancements. An area that has garnered increasing attention is vehicle re-identification—a cutting-edge technology with the primary objective of recognizing the same vehicle across multiple non-overlapping cameras. [1]This field represents a frontier of critical research, driving the development of intelligent traffic systems. With the exceeding popularity as well as widespread usage of Artificial Intelligence and Machine Learning Applications in various fields of businesses, technology, administration, public sectors, creative and applied arts, management etc. [2]The utilization of AI and ML in public transportation has made it more efficient as well as safer. Through deep learning with AI, we are able to solve several issues like:

1. Traffic of vehicles
2. Pollution caused by vehicle emissions,
3. and road accidents.

Through the implementation of vehicle re-identification [3], a robust and efficient surveillance network can be established, greatly benefiting smart cities. This technology plays a pivotal role in optimizing traffic management, enhancing road safety, and fortifying overall security. By automatically detecting, locating, and tracking target vehicles within urban areas [4], it offers immense potential for streamlining traffic flow, bolstering security measures, and facilitating the seamless integration of various smart city applications. As a result, vehicle re-identification has emerged as a cornerstone technology that underpins the evolution of intelligent transportation systems. [5]

The Intelligent Transportation System [6] is a technology in which all vehicles and infrastructure elements, such as road signs and traffic signals, are connected to central control. Its main purpose is to improve flow efficiency, reduce congestion on driveways, prevent accidents, and provide drivers with actual information on road conditions. In Intelligent transportation System, Vehicle-reID is the main component [7] which involves challenging tasks, including dealing with diverse environmental conditions, variations in lighting, and occlusions [8]. Researchers continue to explore and develop novel techniques to enhance the accuracy and robustness of vehicle re-identification systems, making them indispensable components of future smart cities. Vehicle re-ID mainly focuses to identify a vehicle across various surveillance cameras based on their physical appearances. [3]

The system has the ability to inform motorists of hazards and incidents, tell them how fast they are driving, and provide optimal routes. Besides, it can provide better road maintenance information and deliver traffic data to cities for processing and analysis. [9] The survey paper explores various methodologies used in the field of vehicle re-identification (re-ID) to address the challenges of matching and identifying vehicles across multiple cameras and time-frames. It begins by identifying the limitations of traditional vehicle re-ID systems, including issues related to lighting, viewpoints, occlusions, and similar-looking vehicles such as some faced similar result due to light illumination, lack of data or to track spatio-temporal relations. [10] Throughout the survey paper, a comparative analysis of different re-ID approaches are presented against their data sets, features, challenges and strengths, which can help to analyse the various transportation systems.

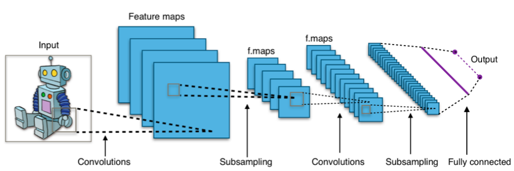

Fig. 1: CNN Architecture

With the help of deep learning models in vehicle re-ID, which leverage global information and reduce distances between identical images while enlarging distances between different ones.We found various approaches for feature extraction, including local feature-based methods, representation learning, and metric learning. [11]Fusion techniques that combine multiple features, such as color, texture, and high-level semantic information, linear SVM (Support Vector Machine) classifier [12], triplet networks etc which all gave very promising results but every notion and experiments had some challenges which are needed to be minimized if not mitigate.

Overall, the survey paper provides a comprehensive overview of the methodologies used in vehicle re-identification research. The paper is organised as follows: the methodology discusses the methods, techniques and systems used to analyse the different approaches , then a comparative study of datasets against their features and results, finally with the results and recommendations of the various papers. We have surveyed approximately fifty papers, which were generally based on vehicle re-identification, most commonly on YOLO algorithm. [13]

2 Methodologies

The subjective papers explored introduces impressive several techniques and methods for vehicle re-identification (re-ID). These techniques aim to address the challenges associated with distinguishing between vehicles with similar appearances and tracking vehicles across different viewpoints.

2.1 A Survey of Vehicle Re-Identification Based on Deep Learning

One approach is a deep learning-based method that divides vehicle re-ID into five categories: local feature-based methods, representation learning, metric learning, unsupervised learning, and attention mechanism. Occlusion is a major challenge in VRI, as it can significantly change the appearance of a vehicle. HONGBO et all discusses a number of methods for dealing with occlusion, such as using context information and using multiple images of the same vehicle. [14]

2.2 A Deep Learning-Based Approach to Progressive Vehicle Re-identification for Urban Surveillance

PROVID (PROgressive Vehicle re-ID) uses the techniques where for the appearance-based coarse filtering, they adopted the fusion model of low-level and high-level features to find the similar vehicles. For license number plate, instead of accurate recognizing the characters of the license plate, they just verified whether two plate images belong to the same vehicle. Therefore, a Siamese neural network was trained with large numbers of plate images for license plate verification. At last, a spatiotemporal relation model is utilized to re-rank vehicles to further improve the final results of vehicle Re-Id. [15]

2.3 Vehicle Re-Identification for Automatic Video Traffic Surveillance

It was observed that some of them utilizes color histograms, histograms of oriented gradients, and a linear regression to solve the re-identification problem. It incorporates supervised learning through a LINEAR SVM CLASSIFIER. The approach used to solve the re-identification problem is by using color histograms and histograms of oriented gradients by a linear regressor. The method attempts to make a 3-D bounding box around the vehicle detected. The features used are: 1.Vector of points describing the 3-D bounding box. 2. color histograms 3. Histogram of oriented gradient features extracted from each vehicle image. 4. One average vector of Color histograms and HOG is also used. The performance of the classifier (in terms of ROC) in natural use would be therefore "close to ideal" and hard to read. This dataset therefore provides a challenging and pessimistic estimation of the classifier performance. Despite that, 60 % of matches can be retrieved (TPR) with only about 10 % of false positives included. Such performance on the difficult dataset promises reliable analysis based on the proposed re-identification method. The main part of the re-identification process is based on regression by a linear classifier. Using this approach, experiment achieved 60 % re-identification true positive rate at 10 % false positives accuracy.The Paper presented a simple and effective solution to the vehicle re-identification problem. The paper proposed a solution which had the problem of noise in the background as they used a 2-D bounding box. They did not use color information of the vehicles which had many false positives. [16]

2.4 A Dual-Path Model With Adaptive Attention For Vehicle Re-Identification

A dual-path adaptive attention model is introduced, consisting of a global appearance path and an orientation conditioned part appearance path. This model enhances the feature learning process for vehicle re-ID.The proposed approach uses a classification method to identify vehicles. Specifically, the model is trained using a combination of triplet loss and cross-entropy loss functions. The triplet loss function is used to learn discriminative features for each vehicle identity, while the cross-entropy loss function is used to classify the orientation of each vehicle into one of eight different classes. For training their data, they used following features as: 1.The global appearance path captures macroscopic features of vehicles, such as color and texture. 2.The orientation conditioned part appearance path focuses on localized discriminative features by attending to informative key-points.The proposed approach uses a classification method to identify vehicles. Specifically, the model is trained using a combination of triplet loss and cross-entropy loss functions. The triplet loss function is used to learn discriminative features for each vehicle identity, while the cross-entropy loss function is used to classify the orientation of each vehicle into one of eight different classes. [17]

2.5 Learning Deep Neural Networks for Vehicle Re-ID with Visual-spatio-temporal Path Proposals

A two-stage framework is proposed, where candidate visual-spatio-temporal paths are suggested in the first stage, and a Siamese-CNN+Path-LSTM network is employed in the second stage to determine the similarity between query pairs for re-ID. The input features used in this approach are: 1. Visual-spatio-temporal states of the vehicle images. 2. Two visual-spatio-temporal states are taken as input. 3. A candidate visual-spatio-temporal path is generated with the two queries as starting and ending states. The network is utilized to determine whether each query pair has the same vehicle identity with spatio-temporal regularization from the candidate path. [18]

2.6 Large Scale Vehicle Re-ID in Urban Surveillance Videos

Six competitive vehicle re-ID methods are implemented as comparative methods, including BOW-SIFT, BOW-CN, AlexNet, GoogleLeNet, AlexNet + BOW-CN, and Fusion of Attributes and Color feaTures (FACT). These methods combine color, texture, and high-level semantic information learned by deep neural networks. The Features used in the proposed method was 1. Texture based feature :- BOW-SIFT 2. Color based feature :- BOW-CN 3. Semantic feature extracted by deep neural network :- AlexNet and GoogleLeNet 4. Feature Fusion :- AlexNet + BOW-CN and Fusion of Attributes and Color feaTures (FACT) . The study solves the problem of domain gap and poor quality of translated images in utilizing synthetic data for object detection, and also focuses on improving training techniques for better performance. [19]

2.7 A Strong Baseline for Vehicle Re-Identification

To bridge the gap between real and synthetic data, the paper suggests using MixStyle Transfer. Additionally, multi-heads with attention mechanisms and the replacement of Triplet Loss with Supervised Contrastive Loss are proposed to enhance performance in vehicle re-ID. [20]

2.8 Part-regularized Near-duplicate Vehicle Re-identification

Region of Information (ROI) is used to capture local information, and the YOLOv4 algorithm along with the SORT tracker is employed to count vehicles accurately.Local features used were: Brands, tags in windows. and non-local features. For part definition: front and back light, front and back window and vehicle brand. They use various classification methods such as k nearest neighbor classification , near-duplicate phenomenon in vehicle re-identification and Local and no-local classification loss. [21]

2.9 Towards Real-time Traffic Flow Estimation using YOLO and SORT from Surveillance Video Footage

They trained YOLOv4 algorithm, SORT tracker to count the vehicles. YOLOv3 showed the best performance in terms of speed and accuracy for their image dataset having a precision value of 0.96 2. SORT is MOT(Multiple Object Detection) algorithm There is a general lack of research in investigating the real-time traffic flow estimation from surveillance video, while also considering movement direction and vehicle class. YOLO is single stage detector and is generally fast and predict the bounding boxes together with classes within a single network pass. Vehicle object detection, classification and tracking are the three main tasks involved while processing video datasets for traffic flow estimation, Authorities can use our algorithm for traffic flow monitoring, traffic anomaly identification, and the development of emergency rescue plans. Also, the responders are able to make management decisions such as detour allocation and changing traffic light timing length during emergencies by using real-time traffic flow. [22]

2.10 Vehicle Re-identification with Viewpoint-aware Metric Learning

Viewpoint-aware Metric Learning is introduced as a novel approach to vehicle re-ID. It learns two metrics for similar and different viewpoints in two feature spaces, respectively, to differentiate between vehicles with nearly identical appearances.The proposed viewpoint-aware metric learning approach learns two separate deep metrics for samples from S-views and D-views, respectively. These metrics are learned using a Siamese network architecture with shared weights and two branches, one for each view. The features are learned by training this network on a large dataset of vehicle images with different viewpoints. [23]

2.11 Vehicle Re-identification in Context (VRIC)

A part-regularized model is proposed to tackle near-duplicate vehicle re-ID, aiming to distinguish between different instances of vehicles with almost identical appearances.The proposed approach in the proposed method uses a discriminative feature preserving method for classification in near-duplicate vehicle re-identification. This method enhances the perceptive ability of subtle discrepancies in appearance between different instances of vehicles. [24]

2.12 Part-regularized Near-duplicate Vehicle Re-identification

The paper presents a Viewpoint-aware Attentive Multi-view Inference (VAMI) model, which leverages visual information to solve the multi-view vehicle re-ID problem. [25]

2.13 Vehicle Re-Identification in Multi-Camera scenarios based on Ensembling Deep Learning Features

The proposed system for vehicle re-identification (ReID) combines different ReID models, including appearance and orientation deep learning features. But it does not mention the specific name of the technique used for vehicle reidentification.deep learning models for feature extraction, including a ResNet50 convolutional neural network pretrained on ImageNet for appearance features extraction . Additionally, the system uses Densenet121 networks with LSR and triplet loss with hard margin for feature extraction and ensemble representation ReID system across multiple cameras based on a feature ensemble representation combining different appearance and structure features. In order to increase the accuracy of the method, it includes a query expansion and temporal pooling of the gallery, followed by the re-ranking of the results and the application of two different methods to add vehicle tracking information. [26]

2.14 A unified neural network for object detection, multiple object tracking and vehicle re-identification

A unified neural network for object detection, multiple object tracking and vehicle re-identification proposed uses End-to-end neural network that combines object detection, multiple object tracking, and vehicle re-identification into a single framework. The network is designed to take in an input image or video stream and output the locations and identities of all objects in the scene, as well as their trajectories over time. The network is trained using a combination of supervised and unsupervised learning techniques, including a triplet loss function for re-identification and a tracklet association module for tracking. It reduces computation while maintaining high accuracy by adding a sibling track branch to the Faster RCNN network and using ROI feature vectors and triplet loss to train this branch. [27]

2.15 Vehicle Re-identification Using Quadruple Directional Deep Learning Features

Vehicle Re-identification Using Quadruple Directional Deep Learning Featureshas proposed a quadruple directional deep learning networks to extract quadruple directional deep learning features (QD-DLF) of vehicle images. The quadruple directional deep learning networks are with similar overall architecture, including the same basic deep learning architecture but different directional feature pooling layers. [28]

2.16 Multiple Object Tracking Using Re-Identification Model with Attention Module

Multiple Object Tracking Using Re-Identification Model with Attention Module used Triplet-based deep learning approach for multi-object tracking. It involves constructing an embedding network with attention modules to enhance performance while preserving computational speed. Also uses an appropriate sampler and mean vectors to address occlusion problems arising from detection-based tracking. CNN is trained to learn a feature embedding that maps each object detection to a high-dimensional feature space. The embedding is designed to be discriminative, meaning that objects belonging to different tracks should be separated in the feature space, while objects belonging to the same track should be close together.The network employs a convolutional attention mechanism to selectively focus on the features of the object, by assigning different weights to different regions. The approach is designed to address occlusion problems arising from detection-based tracking and significantly improves tracking metrics compared to representative tracking systems. [29]

2.17 Orientation Invariant Feature Embedding and Spatial Temporal Regularization for Vehicle Re-identification

Orientation Invariant Feature Embedding and Spatial Temporal Regularization for Vehicle Re-identification proposed framework consists of two main components: the orientation invariant feature embedding component and the spatial-temporal regularization component. [30]

2.18 Unsupervised Vehicle Re-Identification using Triplet Networks

Unsupervised Vehicle Re-Identification using Triplet Networks employs the local constraints and metric learning to accurately predict the identity of a given vehicle. the technique used involves the use of triplet networks. [31]

2.19 Multi-Camera Vehicle Tracking Based on Occlusion-aware and Inter-vehicle Information

It was observed that they used a Kalman filter and Hungarian algorithm for tracking bounding boxes in single-camera vehicle tracking. The proposed framework also incorporates an occlusion-aware module to handle occlusions and suppress false detections. It can be inferred that the method falls under the category of supervised learning. This is because the ReID model is used to extract appearance features from vehicles, which requires labeled training data to learn the mapping between input images and their corresponding features. Also, they used the hierarchical clustering methods to match tracklets. They only consider tracklets from adjacent regions to ensure high-confidence clustering at the first clustering. The remaining trajectories are clustered again to explore matching pairs of tracklets spanning multiple cameras. Finally, they can merge matched tracklets successfully to form completed tracklets. [32]

2.20 CityFlow: A City-Scale Benchmark for Multi-Target Multi-Camera Vehicle Tracking and Re-Identification

It had a very interesting methodology to deploy thereby employed multiple techniques for multi-target multi-camera (MTMC) tracking, including multi-target single-camera (MTSC) tracking, image-based re-identification (ReID), and spatio-temporal data association, also include DeepSORT , TC, and MOANA .They proposes a city-scale benchmark, CityFlow, which enables both video-based MTMC tracking and image-based ReID tasks. The benchmark provides annotations for the original videos, the camera geometry, and calibration information, which can be leveraged to resolve ambiguity in image-based ReID. The proposed benchmark is the first attempt towards city-scale applications in traffic understanding, and has the largest scale among all the existing ReID datasets in terms of spatial coverage and the number of cameras/intersections involved. They presents extensive experiments evaluating the performance of state-of-the-art approaches on the benchmark, comparing and analyzing various visual-spatio-temporal association schemes. These experiments demonstrate that the proposed scenarios are challenging and reflect realistic situations that deployed systems will need to operate in. [33]

2.21 Fast Vehicle Identification via Ranked Semantic Sampling Based Embedding

The proposed method used Ranked Semantic Sampling (RSS) based embedding for efficient cross-view vehicle Re-IDentification.The paper proposes a binary embedding method that preserves the ranked semantic distance between vehicle images from different views. The method uses attribute-guided embedding and manifold learning to discover identity features via a series of comparisons. The proposed method also utilizes a ranked semantic sampling technique to speed up the training process. [34]

2.22 Exploiting Multi-Grain Ranking Constraints for Precisely Searching Visually-similar Vehicles

The Paper proposed to use stochastic gradient descent to conduct the minimization . Additionally, the proposed approaches leverage multi-grain ranking constraints and are implemented with multi-attribute classification in a multi-task deep learning framework The proposed approaches in this paper use a softmax classifier to conduct grain classification . Additionally, the multi-task deep learning framework performs classification tasks on both ID and conventional vehicle attributes such as model and color. [35]

2.23 Deep Relative Distance Learning: Tell the Difference Between Similar Vehicles

This metholodology used the Deep Relative Distance Learning (DRDL) for vehicle re-identification. It is a method that uses a two-branch deep convolutional network to learn the relative distance between two vehicle images and then uses this distance to identify similar vehicles.The DRDL model used extracts deep features from raw vehicle images using a deep convolutional network. These features are then projected into an Euclidean space where the L2 distance can be used to measure the similarity of arbitrary two vehicles. The DRDL model also uses a coupled cluster loss function and a mixed difference network structure to minimize the distances of the same vehicle images and maximize those of other vehicles. [36]

2.24 Heterogeneous Relational Complement for Vehicle Re-identification

Heterogeneous Relational Complement for Vehicle Re-identification have analysed the Heterogeneous Relational Complement Network (HRCN) for vehicle re-identification. The paper proposes to use two types of features for vehicle re-identification: region-specific features and cross-level features. Region-specific features are extracted from specific regions of the vehicle, while cross-level features are extracted from different network stages. [37]

2.25 Self-supervised Geometric Features Discovery via Interpretable Attention for Vehicle Re-Identification and Beyond

The core mechanism of the interpretable attention module used in the Self-supervised Geometric Features Discovery via Interpretable Attention for Vehicle Re-Identification and Beyond framework is to condense the geometric features discovered by the self-supervised learning approach into a compact representation that can be used for vehicle instance discrimination. The proposed framework uses self-supervised learning to discover discriminative geometric features for vehicle re-identification. These features are learned without using accurate supervision and are visualized through attention maps. The attention maps highlight the regions of interest in the vehicle images that are critical for recognition. [38]

2.26 Parsing-based View-aware Embedding Network for Vehicle Re-Identification

Parsing-based View-aware Embedding Network for Vehicle Re-Identification Proposed a parsing-based view-aware embedding network (PVEN) to achieve the view-aware feature alignment and enhancement for vehicle ReID. First, They introduce a parsing network to parse a vehicle into four different views, and then align the features by mask average pooling. Such alignment provides a fine-grained representation of the vehicle. Second, in order to enhance the view-aware features, we design a common-visible attention to focus on the common visible views, which not only shortens the distance among intra-instances, but also enlarges the discrepancy of inter-instances. The PVEN helps capture the stable discriminative information of vehicle under different views. The experiments conducted on three datasets show that our model outperforms state-of-the-art methods by a large margin. [39]

2.27 Robust Wheel Detection for Vehicle Re-Identification

In the paper, the wheel detector constructed in this work uses a retrained SSD network to provide bounding box information for vehicles' wheels. Each detection is accompanied by a confidence score. The proposed approach selects the wheel pair using a vehicle-specific post-processing algorithm that eliminates false detections.the authors utilize deep learning to extract discriminative features from vehicle images. They use a pre-trained object detection model to detect and extract vehicle regions from the input images. Then, they fine-tune the model on a large-scale vehicle dataset to improve its accuracy and efficiency in detecting vehicles. Second, the authors propose a wheel detection module that accurately detects and extracts wheels from vehicle images. The module utilizes a convolutional neural network (CNN) to classify wheel and non-wheel regions in the vehicle image. The authors train the CNN on a large-scale wheel dataset and use it to detect and extract wheels from the input images. Third, the authors utilize the extracted wheel features to re-identify vehicles across different captured images. They compare the wheel features and their associated geometry to match the target vehicle across different images. [40]

2.28 Real-Time Vehicle Orientation Classification and ViewpointAware Vehicle Re-Identificatio

It proposes a new algorithm for vehicle re-identification (re-ID) using a deep learning-based approach with a novel loss function. The proposed algorithm is based on a viewpoint-aware branched network (VABN) architecture that takes into account the viewpoint information of the vehicle images and consists of multiple branches that extract features from different regions of the image. The authors also introduce a new loss function called triplet center loss, which aims to improve the discriminative power of the feature embeddings. [41]

2.29 Unsupervised Vehicle Re-Identification using Triplet Networks

The proposed method employs a metric learning model that learns a representation feature space in which similar vehicles are close together, whereas vehicles with different identities are kept distant. The output vector of the triplet network is used as a feature vector to represent each detected vehicle. The feature vector is generated by comparing the output of the triplet network between different probes and the gallery to generate a ranking. [31]

Overall, these techniques encompass various methodologies, including deep learning, attention mechanisms, metric learning, and multi-view inference, to improve vehicle re-ID performance in areas such as appearance, viewpoint, tracking, and distinguishing between similar instances.

Table 1: Paper and Method Categorization

Category	Method	Characteristics
SUPERVISED	PROgressive Vehicle re-ID (PROVID)	Uses labeled data, license plate verification
	Dual-path adaptive attention model	Uses labeled data, combines global & local features
	Fusion of Attributes and Color Features (FACT)	Uses labeled data, combines color, texture, semantics
	Multi-Scale Vehicle Feature (MSVF) learning model	Uses labeled data, handles scale changes
	Part-regularized model	Uses labeled data, integrates global & local features
	Region Aware deep Model (RAM)	Uses labeled data, combines global & region features
	Viewpoint-aware Attentive Multi-view Inference (VAMI)	Uses labeled data, handles multi-view inference
	Region Batch DropBlock	Uses labeled data, dropout regularization
	Feature Distance Adversarial Network (FDA-Net)	Uses labeled data, generates hard negatives
	End-to-end neural network for detection, tracking, re-identification	Uses labeled data, joint optimization
	Quadruple directional deep learning networks (QD-DLF)	Uses labeled data, directional feature pooling
	Multi-domain learning and Identity Mining	Uses labeled data, domain adaptation
	Adaptive feature learning based on space-time prior	Uses labeled data, adapts features based on prior
	Multi-Camera Vehicle Tracking Based on Occlusion-aware and Inter-vehicle Information	Supervised tracking, occlusion handling
	DeepSORT, TC, MOANA	Supervised learning, spatio-temporal association
	Multi-Grain Ranking Constraints	Supervised, multi-grain similarity ranking
	Deep Relative Distance Learning (DRDL)	Supervised, optimizes relative distance metric
	Heterogeneous Relational Complement Network (HRCN)	Supervised, projects features to new embedding
SEMISUPERVISED	MixStyle Transfer	Uses labeled real & unlabeled synthetic data
	Two-stage framework with Siamese-CNN+Path-LSTM network	Uses labeled & unlabeled data, incorporates spatiotemporal cues
UNSUPERVISED	Unsupervised Vehicle Re-identification using Triplet Networks	Uses unlabeled data, learns feature representation
	Ranked Semantic Sampling (RSS) based embedding	Unsupervised, preserves ranked semantic distance
	Viewpoint-aware Metric Learning	Uses only unlabeled data, learns view-specific metrics
	Self-supervised Geometric Features Discovery via Interpretable Attention	Unsupervised feature learning

3 Databases Used

To solve the Vehicle Re-Identification Problem, several datasets have been used, some of which had upwards of 420,000 images to train machine learning models. Currently as per [42] CPVR-2019 the Veri-Wild data set is the largest Vehicle Reidentification Dataset. The datasets used in most research papers were nicely structured with each image attached with an id label corresponding to its identity in the real world. The images were paired up in the datasets to form two different types of classifications namely - positive samples and negative samples. The positive samples were the ones which had the images of two different instances of the same vehicle to be identified. Whereas the negative samples were the ones which had the images of two different vehicles

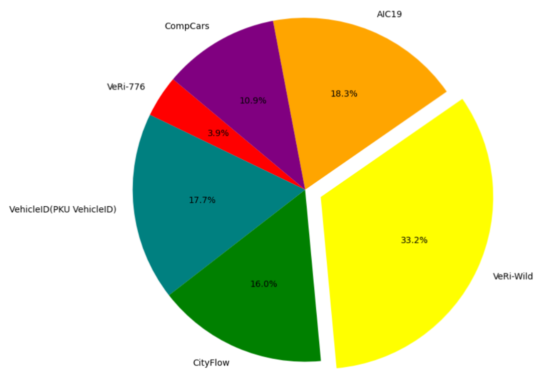

Fig. 2: distribution of images per datasets

which looked similar. The separation of positive samples from the negative ones is what the algorithms are supposed to do to solve the problem of Vehicle Re-identification. Here are some of the most common datasets used in different research papers:-

3.1 VehicleID (PKU VehicleID)

[36]The "VehicleID" dataset contains CARS captured during the daytime by multiple real-world surveillance cameras distributed in a small city in China. There are 26,267 vehicles (221,763 images in total) in the entire dataset. Each image is attached with an id label corresponding to its identity in real world. In addition, the dataset contains manually labelled 10319 vehicles (90196 images in total) of their vehicle model information(i.e."MINI-cooper", "Audi A6L" and "BWM 1 Series"). [43]

Fig. 3: Sample images from VehicleID dataset

3.2 VeRi-776

[44]VeRi-776 is a vehicle re-identification dataset which contains 49,357 images of 776 vehicles from 20 cameras. The dataset is collected in the real traffic scenario, which is close to the setting of CityFlow. The dataset contains bounding boxes, types, colors and brands.The VeRi-776 dataset is divided into two parts: training and testing. The training set contains 39,281 images of 776 vehicles, and the testing set contains 11,076 images of the same 776 vehicles. The images in the training set are labeled with the vehicle's bounding box, type, color, and brand. The images in the testing set are only labeled with the vehicle's bounding box. [43]

Fig. 4: Sample images from VeRi dataset

3.3 CityFlow Dataset

[33] CityFlow is a city-scale traffic camera dataset consisting of more than 3 hours of synchronized HD videos from 40 cameras across 10 intersections, with the longest distance between two simultaneous cameras being 2.5 km. The dataset contains more than 200K annotated bounding boxes covering a wide range of scenes, viewing angles, vehicle models, and urban traffic flow conditions. Camera geometry and calibration information are provided to aid spatio-temporal analysis. In addition, a subset of the benchmark is made available for the task of image-based vehicle re-identification (ReID). [45]

3.4 AIC19 Dataset

[46] AIC19 vehicle tracking dataset, which contains 36 labeled videos for training. The authors construct their training dataset by concatenating neighboring frames in the divided AIC19 train dataset, resulting in 17,556 images and 683 vehicle instances. The validation dataset consists of 3 videos and contains 2,988 images and 97 vehicle instances. [47]

3.5 Veri-Wild

[48] Veri-Wild is the largest vehicle re-identification dataset (as of CVPR 2019) [42]. The dataset is captured from a large CCTV surveillance system consisting of 174 cameras across one month (30 24h) under unconstrained scenarios. This dataset comprises 416,314 vehicle images of 40,671 identities. Evaluation on this dataset is split across three subsets: small, medium and large; comprising 3000, 5000 and 10,000 identities respectively (in probe and gallery sets). [39]

3.6 CompCars Dataset

The Comprehensive Cars (CompCars) dataset [49] contains data from two scenarios, including images from web-nature and surveillance-nature. The web-nature data contains 163 car makes with 1,716 car models. There are a total of 136,726 images capturing the entire cars and 27,618 images capturing the car parts. The full car images are labeled with bounding boxes and viewpoints. Each car model is labeled with five attributes, including maximum speed, displacement, number of doors, number of seats, and type of car. The surveillance-nature data contains 50,000 car images captured in the front view. Please refer to the paper for the details. [49] The dataset is well prepared for the following computer vision tasks: Fine-grained classification Attribute prediction Car model verification The train/test subsets of these tasks introduced in our paper are included in the dataset. Researchers are also welcome to use it for any other tasks such as image ranking, multi-task learning, and 3D reconstruction. [34]

4 Results and Discussion

4.1 Overview

Vehicle re-identification matches vehicles across different cameras. This can be challenging because vehicles often look similar, changes in lighting and cameras, and getting blocked by other objects. [50] Most of the old methods struggled with this in uncontrolled areas like cities. Recent studies combine a vehicle's overall appearance with local details like wheels, plates, and shape. Others used how vehicles moved over time and some generated computer vehicle images to expand training data. The new techniques were tested on some standard datasets. [48] Overall, they worked much better than older simple methods. But matching vehicles across huge camera networks in uncontrolled places is still tricky. These advancements could lead to practical implementations such as tracking vehicles between cameras, conducting comprehensive searches across multiple cameras for a specific vehicle, and studying vehicle behavior. To build useful systems, the methods need to handle more cameras and vehicles. And they should merge vehicle re-identification with detecting, tracking, describing, and predicting vehicles. Vehicle re-id is crucial for intelligent transportation. The main contributions seen in recent papers were: Techniques that combined how vehicles looked overall with local features and how they moved. Some used computer-generated vehicles to handle limited real data. Tests on datasets show big gains over the previous best methods. But scaling to large uncontrolled environments remains hard. Ideas for upcoming work. This includes combining learning across computer vision areas and building useful systems for smart cities. Overall, vehicle re-identification has advanced a lot recently. As progress continues, it could enable the next generation of intelligent transportation technology. [51]

4.2 Evaluation Strategy

The researchers came up with some ways to measure how well the vehicle re-identification model works. This section will talk about a few common ways to measure how well the vehicle re-identification algorithm works.

4.2.1 Rank

To use the rank method, the different methods or approaches are first evaluated against the criteria. Then, they are ranked from best to worst. The method or approach with the highest rank is considered to be the best. The rank-n value is the probability that the correct vehicle is in the top n positions of the search results. A higher rank suggests that the model is performing better. [52]

4.2.2 CMC Curve

The CMC curve [53] shows the probability of a correct match occurring within the top-K candidate matches returned by the algorithm. It plots the identification rate as a function of the rank position K. For example, the CMC rank-1 score indicates the percentage of times the correct match was found in the top 1 result. A CMC rank-5 score shows the rate for finding the correct match within the top 5 results. Higher CMC scores at lower rank values indicate better performance. The CMC curve provides insight into how well the algorithm can return the correct match within tighter candidate sets. Steeper CMC curves that quickly approach 100% demonstrate excellent retrieval and ranking capabilities. The area under the CMC curve summarizes the overall algorithm accuracy. [54]

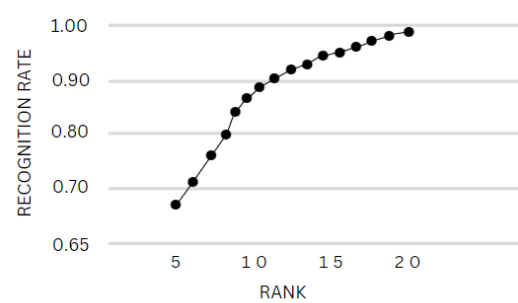

Fig. 5: An example of a CMC Curve

4.2.3 mAP

Mean average precision (mAP) [55] is a metric that measures how well a model judges the results on all query images. It is calculated by averaging the APs for all of the query images. A higher mAP indicates that the model is better at judging the results on all query images. mAP is a useful metric for evaluating the performance of object detection models. It is a comprehensive metric that takes into account both precision and recall. Additionally, mAP is a scale-invariant metric, which means that it is not affected by the number of objects in the image.

4.3 Results

The outcomes of the approaches demonstrated in Table 1 have shown notable improvements in vehicle re-identification accuracy, as evident from increased mean average precision

Table 2: Proposed Methods and Results in Vehicle Re-Identification Research

Papers

Dual-path adaptive attention model	VERI	Rank-1 Accuracy: 94.5%, Mean Average Precision (mAP): 87.7%
Two-stage framework with Siamese-CNN+Path-LSTM network	VERI	Mean Average Precision (mAP): 58.27%
Fusion of attributes and color features (FACT)	VERI	mAP achieved by FACT is 58.27%
MixStyle Transfer for real-synthetic data gap bridging	CityFlow and VeRi	Baseline mAP with Cross Entropy Loss weight: 73.3%, Baseline mAP with Triplet Loss weight and MALW: 78.4%
Viewpoint-aware Metric Learning	VERI and VehicleID	VehicleID Dataset: Top-1 Accuracy on Small subset: 96.03%, Top-1 Accuracy on Medium subset: 91.98%, Top-1 Accuracy on Large subset: 87.68%, Veri-776 Dataset: Rank-1 Accuracy: 82.17%
Multi-Scale Vehicle Feature (MSVF) learning model	VehicleID, VeRi-776	VehicleID: Rank 1 accuracy: 63.02%, VeRi: Rank 1 accuracy: 88.56%
Part-regularized model for near-duplicate vehicle re-identification	VehicleID, VeRi-776	VehicleID: Rank-1 accuracy of 94.5% and a mAP of 84.5%, VeRi: Rank-1 accuracy of 94.3% and a mAP of 74.3%
Region Aware deep Model (RAM)	VeRi and VehicleID	VeRi: Improvement over baseline feature: 6.5% in mAP, achieves the best Rank-1 and Rank-5 accuracies in vehicle re-identification
Viewpoint-aware Attentive Multi-view Inference (VAMI)	VeRi and VehicleID	VeRi: Multi-view feature space gain: 8.19% in mAP, achieves the best performance with a 2.13% increase in mAP compared to the state-of-the-art methods
Region Batch DropBlock	VehicleID and Veri776	VeRi: mAP- 81.9, CMC@1-96.3, VehicleID(Small): CMC@1-85.5, CMC@5-97.0
Feature Distance Adversarial Network (FDA-Net)	VERI-Wild	FDA-Net achieves the highest Rank-1 accuracy: 55.49%, Previous state-of-the-art method Rank-1 accuracy: 47.85%.
End-to-end neural network for object detection, tracking, and re-identification	AIC19	AIC19 Vehicle Tracking Dataset: mAP: 57.79% (significant improvement over the previous method), MOTA: 47.5%.
Self-supervised Geometric Features Discovery via Interpretable Attention	VeRi-776 and CityFlow-ReID	VeRi-776 Dataset: achieves a Rank-1 accuracy of 94.5%, outperforms other state-of-the-art methods on the VeRi-776 dataset, CityFlow-ReID Dataset: achieves approximately 7.0% improvement in mAP compared to the second-best method, achieves approximately 6.0% improvement in Top-1 accuracy compared to the second-best method, achieves approximately 1.0% improvement in Top-5 accuracy compared to the second-best method, outperforms other state-of-the-art methods
VehicleNet: Two-Stage Progressive Learning Approach	AICity Challenge private test set, VeRi-776, VehicleID, and CompCar	AICity Challenge Private Test Set: VehicleNet achieves a state-of-the-art accuracy of 86.07% mAP, The proposed two-stage progressive approach outperforms other methods on the AICity Challenge Private Test Set. VeRi-776 Dataset: VehicleNet achieves competitive results on the VeRi-776 dataset, VehicleID Dataset: VehicleNet achieves competitive results on the VehicleID dataset.
Parsing-based View-aware Embedding Network (PVEN)	VeRi-776, VehicleID, and VERI-Wild	VeRi-776 Dataset: mAP on the small dataset: 87.8%, mAP on the medium dataset: 84.8%, mAP on the large dataset: 68.8%, The proposed PVEN method outperforms previous state-of-the-art methods by a large margin, with improvements of 47.4%, 47.2%, and 46.9% on the small, medium, and large datasets, respectively. CMC@1 and CMC@5 performance is also significantly better than the previous state-of-the-art methods.
Robust Wheel Detection for Vehicle Re-Identification	VehicleID	Improved accuracy in vehicle re-identification, The proposed approach achieved promising results on the VehicleID dataset, outperforming several state-of-the-art methods with a rank-1 accuracy of 94.5% and an mAP of 87.5%
Real-Time Vehicle Orientation Classification and Viewpoint-Aware Vehicle Re-Identification	VeRi-776 and VehicleID	Improved accuracy in vehicle re-identification, The proposed viewpoint-aware multi-branch network improved the performance of vehicle re-ID, achieving an increase in mAP and rank-1 score compared to the ResNet-50 baseline, State-of-the-art performance on VeRi-776 and VehicleID datasets
Quadruple directional deep learning networks (QD-DLF)	VeRi and VehicleID	VeRi-776 Dataset: achieves the highest MAP of 61.83%, achieves the highest Rank-1 identification rate of 88.50% among all methods compared, outperforms multiple state-of-the-art methods on the VeRi dataset. VehicleID Dataset: outperforms all deep learning-based methods under comparison, including DJDL, DenseNet121, Improved Triplet CNN, DRDL, FACT, NuFACT, and GoogLeNet.
Multi-domain learning and Identity Mining	CityFlow	Achieves a state-of-the-art mAP score of 59.7%
Adaptive feature learning based on space-time prior	VeRi CompCars BoxCars	Increased mean average precision (mAP) and higher rank-1 accuracy
Multi-Camera Vehicle Tracking Based on Occlusion-aware and Inter-vehicle Information	CityFlowV2 dataset	Improved accuracy in vehicle re-identification (IDF1 improvement from 79.58% to 82.85%) and outperforming other methods in terms of mAP (IDF1), top-1 (IDP), and top-5 (IDR) accuracies.
DeepSORT, TC, MOANA	CityFlow dataset	Increased mean average precision (mAP), Outperforming other methods in terms of mAP
Ranked Semantic Sampling (RSS) based embedding	CompCars and VeRi	CompCars Dataset: achieves a Rank-1 accuracy of 94.6%, achieves an AUC (Area Under the Curve) score of 98.8%, outperforms several state-of-the-art methods. VeRi Dataset: achieves a Rank-1 accuracy of 79.7%, achieves an AUC score of 94.7%, outperforms several state-of-the-art methods.
Deep Relative Distance Learning (DRDL)	VehicleID and CompCars	VehicleID Dataset: achieves a Rank-1 accuracy of 94.5%, achieves an mAP of 87.5%, achieves an NDCG (Normalized Discounted Cumulative Gain) of 93.2%, outperforms all other methods in terms of Rank-1 accuracy, mAP, and NDCG. CompCars Dataset: achieves a Rank-1 accuracy of 87.5%, achieves an mAP of 75.3%, achieves an NDCG of 85.2%, outperforms all other methods in terms of Rank-1 accuracy, mAP, and NDCG.
Heterogeneous Relational Complement Network (HRCN)	VeRi-776, VehicleID, and VERI-Wild	VeRi-776 Dataset: mCGM: 0.630, mAP: 0.831, CMC@5: 0.989, VehicleID Dataset: mAP: 0.868, VERI-Wild Dataset: mAP for subsets with 3,000 identities: 0.536, mAP for subsets with 5,000 identities: 0.464, mAP for subsets with 10,000 identities: 0.394

(mAP), higher rank-1 accuracy, and outperforming state-of-the-art methods. They establish new benchmarks on prominent datasets like VehicleID, VeRi-776, CityFlow, AIC19, and Veri-Wild.

The practical implications of these advancements are significant, including applications in traffic flow monitoring, traffic anomaly detection, and emergency response planning. The proposed algorithms enable real-time analysis of traffic flow and support decision-making in traffic management systems.

4.4 Possible Future Outcomes

Vehicle re-identification is a core technology in intelligent transportation and monitoring. Deep learning has been widely used in vehicle re-identification in recent years. Based on our survey, we present several future research directions, including:

•Assistance of spatiotemporal information: Spatiotemporal information [56] can be used to improve the performance of vehicle re-identification. For example, we can use the near-to-far principle to search for vehicles in a time and space range. However, it is still challenging to effectively use spatiotemporal information in real-world scenarios.

•Datasets with more information: Existing vehicle re-identification datasets do not provide all the necessary information, such as original video and camera correction information. This limits the development of vehicle re-identification methods.

•Multi-view re-identification: Vehicle re-identification [57] is more challenging than person re-identification because of the high variability within the class and the high similarity between the classes. In the future, we can use GANs to learn the correlation between different perspectives and synthesize multi-view features.

•Combination of detection and re-identification tasks: Current vehicle re-identification methods assume that the boundary box of the vehicle is accurate. However, this is not always the case. In the future, we can combine the detection and re-identification tasks to improve the accuracy of vehicle re-identification.

•Integration of multiple approaches: Different approaches have their own advantages and disadvantages. In the future, we can integrate multiple approaches to improve the accuracy of vehicle re-identification. Application of transfer learning to real traffic scenes: Transfer learning allows us to apply vehicle re-identification models to real traffic scenes even if the training data is different from the testing data.

5 Conclusion

In this survey, we extensively reviewed research papers on deep learning-based [58] vehicle re-identification (V-reID). V-reID has become crucial for public safety and smart city development due to the growing number of vehicles and the need for surveillance footage analysis. Our analysis compared these methods, highlighting their characteristics, advantages, and disadvantages. These include variations in vehicle appearance, occlusions, lighting conditions, and scalability to large-scale surveillance scenarios. We have also highlighted open issues and discussed potential research directions for future investigations. We also summarized recent vehicle datasets, discussed evaluation strategies, and identified challenges and potential research directions. This survey provides a valuable resource for researchers, aiding in the selection of appropriate techniques for V-reID applications and emphasizing the need for large-scale real surveillance datasets to improve model performance.

@articlesun2021tbe, title=TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification, author=Sun, Wei and Dai, Guangzhao and Zhang, Xiaorui and He, Xiaozheng and Chen, Xuan, journal=IEEE Transactions on Intelligent Transportation Systems, volume=23, number=9, pages=14557–14569, year=2021, publisher=IEEE

References

[1] Momin, B. F., and Mujawar, T. M., 2015. "Vehicle detection and attribute based search of vehicles in video surveillance system". In 2015 international conference on circuits, power and computing technologies [ICCPCT-2015], IEEE, pp. 1–4.

[2] Wang, F.-Y., 2010. "Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications". *IEEE Transactions on Intelligent Transportation Systems,* **11**(3), Sep., pp. 630–638.

[3] Khan, S. D., and Ullah, H., 2019. "A survey of advances in vision-based vehicle re-identification". *Computer Vision and Image Understanding,* **182**, pp. 50–63.

[4] He, Z., Lei, Y., Bai, S., and Wu, W., 2019. "Multi-camera vehicle tracking with powerful visual features and spatial-temporal cue.". In CVPR Workshops, Vol. 1.

[5] Wang, H., Hou, J., and Chen, N., 2019. "A survey of vehicle re-identification based on deep learning". *IEEE Access,* **7**, pp. 172443–172469.

[6] Figueiredo, L., Jesus, I., Machado, J., Ferreira, J., and Martins de Carvalho, J., 2001. "Towards the development of intelligent transportation systems". In ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585), pp. 1206–1211.

[7] Guo, H., Zhu, K., Tang, M., and Wang, J., 2019. "Two-level attention network with multi-grain ranking loss for vehicle re-identification". *IEEE Transactions on Image Processing,* **28**(9), pp. 4328–4338.

[8] Zakria, Cai, J., Deng, J., Aftab, M. U., Khokhar, M. S., and Kumar, R., 2019. "Efficient and deep vehicle re-identification using multi-level feature extraction". *Applied Sciences,* **9**(7), p. 1291.

[9] Wu, Z., Xu, T., Wang, F., Wang, X., and Song, J., 2023. "Complete solution for vehicle re-id in surround-view camera system". *International Journal of Machine Learning and Cybernetics,* **14**(5), pp. 1739–1749.

[10] Li, P., Li, G., Yan, Z., Li, Y., Lu, M., Xu, P., Gu, Y., Bai, B., Zhang, Y., and Chuxing, D., 2019. "Spatio-temporal consistency and hierarchical matching for multi-target multi-camera vehicle tracking.". In CVPR Workshops, pp. 222–230.

[11] Mathews, M., and Sethilnathan, T., 2023. "A study of cnn models for re-identification of vehicles". In 2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), IEEE, pp. 01–08.

[12] Keerthi, S. S., Shevade, S. K., Bhattacharyya, C., and Murthy, K. R. K., 2001. "Improvements to platt's smo algorithm for svm classifier design". *Neural computation,* **13**(3), pp. 637–649.

[13] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., 2016. "You only look once: Unified, real-time object detection". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779–788.

[14] Wang, H., Hou, J., and Chen, N., 2019. "A survey of vehicle re-identification based on deep learning". Vol. 7, IEEE, pp. 172443–172469.

[15] Liu, X., Liu, W., Mei, T., and Ma, H., 2016. "A deep learning-based approach to progressive vehicle re-identification for urban surveillance". In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14, Springer, pp. 869–884.

[16] Zapletal, D., and Herout, A., 2016. "Vehicle re-identification for automatic video traffic surveillance". In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 25–31.

[17] Khorramshahi, P., Kumar, A., Peri, N., Rambhatla, S. S., Chen, J.-C., and Chellappa, R., 2019. "A dual-path model with adaptive attention for vehicle re-identification". In Proceedings of the IEEE/CVF international conference on computer vision, pp. 6132–6141.

[18] Shen, Y., Xiao, T., Li, H., Yi, S., and Wang, X., 2017. "Learning deep neural networks for vehicle re-id with visual-spatio-temporal path proposals". In Proceedings of the IEEE international conference on computer vision, pp. 1900–1909.

[19] Liu, X., Liu, W., Ma, H., and Fu, H., 2016. "Large-scale vehicle re-identification in urban surveillance videos". In 2016 IEEE international conference on multimedia and expo (ICME), IEEE, pp. 1–6.

[20] Huynh, S. V., 2021. "A strong baseline for vehicle re-identification". In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4147–4154.

[21] He, B., Li, J., Zhao, Y., and Tian, Y., 2019. "Part-regularized near-duplicate vehicle re-identification". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3997–4005.

[22] Algiriyage, N., Prasanna, R., Doyle, E., Stock, K., Johnston, D., Punchihewa, M., and Jayawardhana, S., 2021. "Towards real-time traffic flow estimation using yolo and sort from surveillance video footage". In Proceedings of the 18th International Conference on Information Systems for Crisis Response and Management, Blacksburg, VA, USA, pp. 23–26.

[23] Chu, R., Sun, Y., Li, Y., Liu, Z., Zhang, C., and Wei, Y., 2019. "Vehicle re-identification with viewpoint-aware metric learning". In Proceedings of the IEEE/CVF international conference on computer vision, pp. 8282–8291.

[24] Kanacı, A., Zhu, X., and Gong, S., 2019. "Vehicle re-identification in context". In Pattern Recognition: 40th German Conference, GCPR 2018, Stuttgart, Germany, October 9-12, 2018, Proceedings 40, Springer, pp. 377–390.

[25] He, B., Li, J., Zhao, Y., and Tian, Y., 2019. "Part-regularized near-duplicate vehicle re-identification". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3997–4005.

[26] Moral, P., Garcia-Martin, A., and Martinez, J. M., 2020. "Vehicle re-identification in multi-camera scenarios based on ensembling deep learning features". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 604–605.

[27] Xu, Y., and Wang, J., 2019. "A unified neural network for object detection, multiple object tracking and vehicle re-identification". *arXiv preprint arXiv:1907.03465*.

[28] Zhu, J., Zeng, H., Huang, J., Liao, S., Lei, Z., Cai, C., and Zheng, L., 2019. "Vehicle re-identification using quadruple directional deep learning features". *IEEE Transactions on Intelligent Transportation Systems,* **21**(1), pp. 410–420.

[29] Ahn, W.-J., Ko, K.-S., Lim, M.-T., Pae, D.-S., and Kang, T.-K., 2023. "Multiple object tracking using re-identification model with attention module". *Applied Sciences,* **13**(7), p. 4298.

[30] Wang, Z., Tang, L., Liu, X., Yao, Z., Yi, S., Shao, J., Yan, J., Wang, S., Li, H., and Wang, X., 2017. "Orientation invariant feature embedding and spatial temporal regularization for vehicle re-identification". In Proceedings of the IEEE international conference on computer vision, pp. 379–387.

[31] Antonio Marin-Reyes, P., Palazzi, A., Bergamini, L., Calderara, S., Lorenzo-Navarro, J., and Cucchiara, R., 2018. "Unsupervised vehicle re-identification using triplet networks". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 166–171.

[32] Liu, Y., Zhang, X., Zhang, B., Zhang, X., Wang, S., and Xu, J., 2022. "Multi-camera vehicle tracking based on occlusion-aware and inter-vehicle information". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3257–3264.

[33] Tang, Z., Naphade, M., Liu, M.-Y., Yang, X., Birchfield, S., Wang, S., Kumar, R., Anastasiu, D., and Hwang, J.-N., 2019. "Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8797–8806.

[34] Zheng, F., Miao, X., and Huang, H., 2018. "Fast vehicle identification in surveillance via ranked semantic sampling based embedding". In 27th International Joint Conference on Artificial Intelligence (IJCAI 2018).

[35] Yan, K., Tian, Y., Wang, Y., Zeng, W., and Huang, T., 2017. "Exploiting multi-grain ranking constraints for precisely searching visually-similar vehicles". In Proceedings of the IEEE international conference on computer vision, pp. 562–570.

[36] Liu, H., Tian, Y., Yang, Y., Pang, L., and Huang, T., 2016. "Deep relative distance learning: Tell the difference between similar vehicles". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2167–2175.

[37] Zhao, J., Zhao, Y., Li, J., Yan, K., and Tian, Y., 2021. "Heterogeneous relational complement for vehicle re-identification". In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 205–214.

[38] Li, M., Huang, X., and Zhang, Z., 2021. "Self-supervised geometric features discovery via interpretable attention for vehicle re-identification and beyond". In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 194–204.

[39] Meng, D., Li, L., Liu, X., Li, Y., Yang, S., Zha, Z.-J., Gao, X., Wang, S., and Huang, Q., 2020. "Parsing-based view-aware embedding network for vehicle re-identification". In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 7103–7112.

[40] Ghanem, S., and Kerekes, R. A., 2022. "Robust wheel detection for vehicle re-identification". *Sensors,* **23**(1), p. 393.

[41] Tamas Kocsis, O., Alkanat, T., Bondarev, E., et al., 2021. "Real-time vehicle orientation classification and viewpoint-aware vehicle re-identification". *Electronic Imaging,* **2021**(10), pp. 234–1.

[42] Zhedong Zheng, M. C., 2022. Vehicle re-id collection.

[43] He, B., Li, J., Zhao, Y., and Tian, Y., 2019. "Part-regularized near-duplicate vehicle re-identification". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3997–4005.

[44] Liu, X., Liu, W., Ma, H., and Fu, H., 2016. "Large-scale

vehicle re-identification in urban surveillance videos". In 2016 IEEE international conference on multimedia and expo (ICME), IEEE, pp. 1–6.

[45] He, S., Luo, H., Chen, W., Zhang, M., Zhang, Y., Wang, F., Li, H., and Jiang, W., 2020. "Multi-domain learning and identity mining for vehicle re-identification". In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 582–583.

[46] Naphade, M., Wang, S., Anastasiu, D. C., Tang, Z., Chang, M.-C., Yao, Y., Zheng, L., Rahman, M. S., Arya, M. S., Sharma, A., Feng, Q., Ablavsky, V., Sclaroff, S., Chakraborty, P., Prajapati, S., Li, A., Li, S., Kunadharaju, K., Jiang, S., and Chellappa, R., 2023. "The 7th ai city challenge". In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops.

[47] Chang, M.-C., Chiang, C.-K., Tsai, C.-M., Chang, Y.-K., Chiang, H.-L., Wang, Y.-A., Chang, S.-Y., Li, Y.-L., Tsai, M.-S., and Tseng, H.-Y., 2020. "Ai city challenge 2020-computer vision for smart transportation applications". In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp. 620–621.

[48] Lou, Y., Bai, Y., Liu, J., Wang, S., and Duan, L.-Y., 2019. "Veri-wild: A large dataset and a new method for vehicle re-identification in the wild". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3235–3243.

[49] Yang, L., Luo, P., Change Loy, C., and Tang, X., 2015. "A large-scale car dataset for fine-grained categorization and verification". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3973–3981.

[50] Davies, S., 2022. "What is occlusion?". In *A Guide to Good Occlusal Practice*. Springer, pp. 1–22.

[51] Hsu, H.-M., Wang, Y., and Hwang, J.-N., 2020. "Traffic-aware multi-camera tracking of vehicles based on reid and camera link model". In Proceedings of the 28th ACM International Conference on Multimedia, pp. 964–972.

[52] Wilcoxon, F., 1992. "Individual comparisons by ranking methods". In *Breakthroughs in Statistics: Methodology and Distribution*. Springer, pp. 196–202.

[53] Damer, N., Opel, A., and Nouak, A., 2014. "Cmc curve properties and biometric source weighting in multi-biometric score-level fusion". In 17th International Conference on Information Fusion (FUSION), IEEE, pp. 1–6.

[54] Perdomo, O., 2012. "A dynamical interpretation of the profile curve of cmc twizzler surfaces". *Pacific Journal of Mathematics,* **258**(2), pp. 459–485.

[55] Henderson, P., and Ferrari, V., 2017. "End-to-end training of object class detectors for mean average precision". In Computer Vision–ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part V 13, Springer, pp. 198–213.

[56] Chen, S.-C., Shyu, M.-L., and Zhang, C., 2001. "An intelligent framework for spatio-temporal vehicle tracking". In ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 01TH8585), IEEE, pp. 213–218.

[57] Zhou, Y., and Shao, L., 2018. "Aware attentive multi-view inference for vehicle re-identification". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6489–6498.

[58] LeCun, Y., Bengio, Y., and Hinton, G., 2015. "Deep learning". *nature,* **521**(7553), pp. 436–444.