AML Final Project Report

Title: Vehicle re identification in Deep Learning

Domain: Computer Vision

Student Name: Lava Kumar Saripudi **Instructor: Dr. CJ Wu**

Abstract:

The state-of-the-art deep learning models and methods for vehicle re-identification are briefly summarized in this report. It provides an overview of the most recent methods, their efficacy, and the difficulties encountered in this area. Though there are many methods used for vehicle re identification. This report focuses on four important publications that make use of transformers, co-segmentation, Siamese networks, and orientation-aware techniques. Metrics like rank-1 accuracy and mAP are used to assess the efficacy of these strategies. In addition, the study examines the existing and future industrial uses of vehicle re-identification in security, transportation, and healthcare. Finally, it identifies the shortcomings of deep learning models and suggests potential workarounds. This study advances vehicle re-identification technology and its use in real-world situations.

Acknowledgements:

I want to express my gratitude to everyone who reviewed, particularly the students, technical support personnel, and reviewers. I want to give thanks to:

- Dr. Chaojiang (CJ) Wu, Ph.D.
- State-of-art Techniques Team
- Vehicle Re Identification Competitors

Table Of Contents:

1.Introduction:

Vehicle Re Identification in Deep Learning:

Re-identifying vehicles is a crucial computer vision problem that seeks to recognize and track vehicles across various cameras or video frames. In this area, deep learning models have shown findings that are encouraging, offering precise and effective solutions. The development of efficient models and algorithms to recognize and track cars across several cameras and time frames is the main goal of research on vehicle re-identification in deep learning. Convolutional neural networks (CNNs), Siamese networks, and transformer-based models are some examples of deep learning techniques that are used to extract discriminative features from car photos and build reliable representations for precise identification. Deep learning techniques for vehicle re-identification have great promise for several sectors, including surveillance and security, transportation administration, the auto industry, and urban planning. Increased safety precautions, greater urban mobility, and further development of intelligent transportation systems will all be facilitated by continued research and development in this area. Although vehicle re-identification based on deep learning has demonstrated encouraging results, there are still certain difficulties and restrictions. These include occlusions, complex backdrops, varying lighting and weather, and a lack of labeled data for training. Advancements in data gathering, augmentation methods, and reliable models that can handle real-world events are all necessary to address these difficulties.

Scope:

The goal of research on deep learning models and algorithms for vehicle re-identification is to increase the reliability and accuracy of vehicle identification systems. The goal of the project is to create methods that can successfully match and identify automobiles in a variety of camera angles, lighting situations, occlusions, and image alterations. Examining the possibility of deep learning models, such as Siamese networks, transformer-based models, and co-segmentation techniques, to identify and learn distinguishing features from vehicle photos is also included in the scope. The focus also includes analyzing present difficulties and constraints in vehicle re-identification, such as the need for large, labeled datasets and robustness to environmental fluctuations, and exploring potential solutions to these constraints.

State-Of-the-Art Techniques list in Vehicle Re identification:

Rank	Model	Rank-1	Rank-5	Rank1	Rank5	mAP	Extra Training Data	Paper	Code	Result	Year	Tags
1	**Recall@k Surrogate loss** (ViT-B/16)	94.7	97.1				×	Recall@k Surrogate Loss with Large Batches and Similarity Mixup			2021	
2	**Recall@k Surrogate loss** (ResNet-50)	93.8	96.6				×	Recall@k Surrogate Loss with Large Batches and Similarity Mixup			2021	512 ResNet-50
3	**PNP Loss**	93.2	96.6				×	Rethinking the Optimization of Average Precision: Only Penalizing Negative Instances before Positive Ones is Enough			2021	
4	**RPTM**	92.9	96.3			80.5	×	Relation Preserving Triplet Mining for Stabilising the Triplet Loss in Re-identification Systems			2021	
5	**Smooth-AP**	91.9	96.2				×	Smooth-AP: Smoothing the Path Towards Large-Scale Image Retrieval			2020	
6	**ANet**	80.5	94.6	80.5	94.6		×	AttributeNet: Attribute Enhanced Vehicle Re-Identification			2021	
7	**vehiclenet**	79.46					✓	VehicleNet: Learning Robust Feature Representation for Vehicle Re-identification			2020	
8	**MSINet** (2.3M w/o RK)	77.9	91.7				×	MSINet: Twins Contrastive Search of Multi-Scale Interaction for Object ReID			2023	

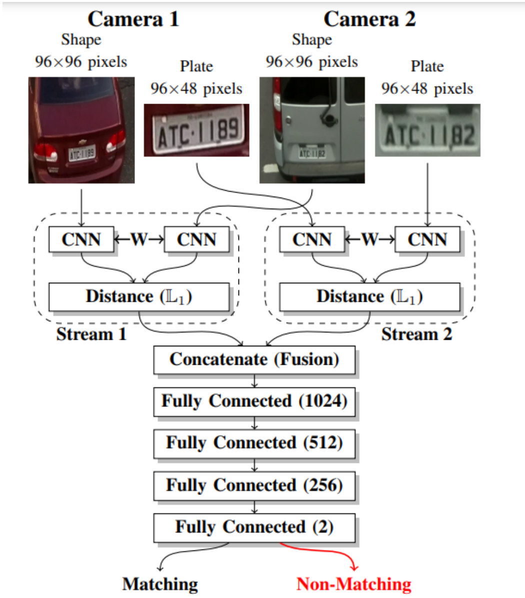
2.Literature Review:

Re-identifying vehicles with the intention of matching them across various cameras or instances is a difficult job in the field of computer vision. The accuracy and robustness of vehicle re-identification have seen substantial improvements thanks to deep learning models and algorithms. In-depth study of current research in vehicle re-identification is provided in this literature review, with an emphasis on cutting-edge methods, their technical specifics, efficacy, and a thorough comparison of various approaches.

A two-stream Siamese neural network for vehicle re-ID is the first method. To capture both fine-grained visual details and spatial interactions, the model combines appearance and spatial stream networks. While the spatial stream concentrates on extracting spatial context information using a spatial transformer network, the appearance stream uses a CNN-based architecture to extract discriminative appearance features. With the help of a fusion module, the two streams are mixed. Extensive tests on benchmark datasets show that the suggested approach is effective, reaching state-of-the-art vehicle re-ID performance.

The Siamese neural network design used in this method has two streams, each of which contains a convolutional neural network (CNN). The two streams individually process the query and gallery photos, removing visual elements from each picture. The matching score is then calculated by comparing these features using a similarity metric. This method has proven to be quite good at handling different viewpoints, lighting conditions, and occlusions while still accurately re-identifying the vehicle.

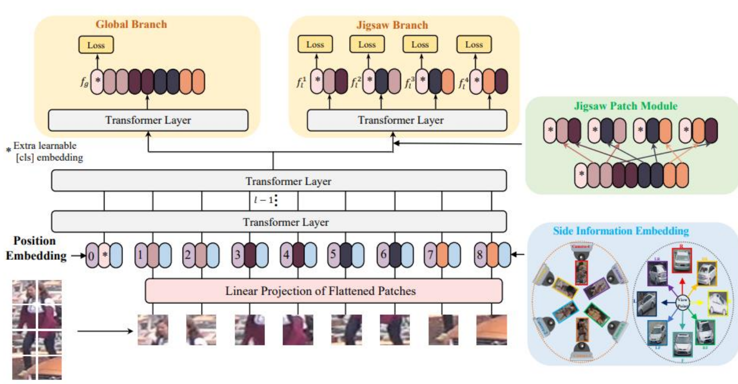
Flowchart of the proposed Two-Stream Siamese for Vehicle Matching's inference process



The second approach TransReID makes use of transformer-based models, which were initially developed for natural language processing, for object re-ID, including vehicle re-ID. The supplied vehicle photos are transformed into long-range dependencies and global contextual information. To accurately match across variations in viewpoint, lighting, and occlusions, it uses self-attention mechanisms to understand correlations between various vehicle components.

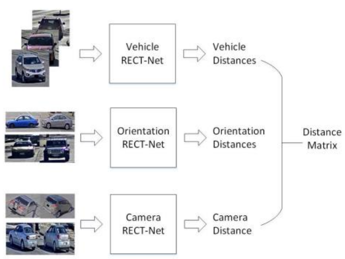
The performance of this method in vehicle re-identification tasks has proven to be cutting-edge. The excellent feature extraction and context modeling made possible by the transformer design improve the discriminative ability of the learned representations. The new TransReID model outperforms conventional CNN-based approaches [1,2,3,4,5]., as shown by experimental results on vehicle re-ID datasets, which show significant performance gains.

Framework of proposed TransReID



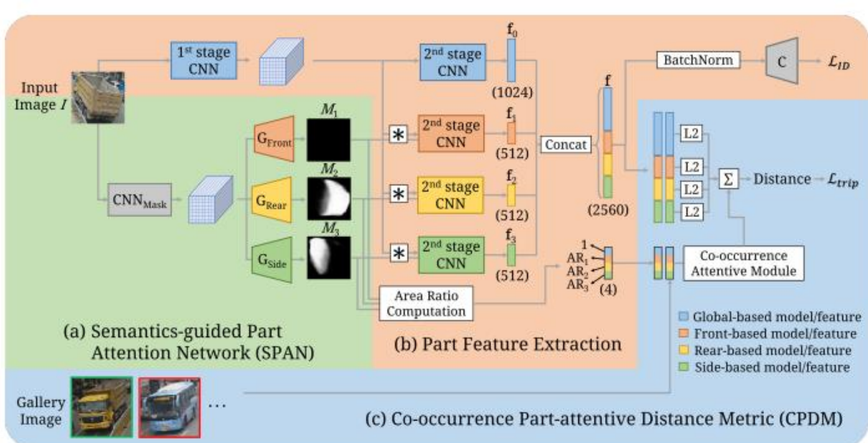
The third method, called VOC-ReID, incorporates vehicle-object co-segmentation methods into the pipeline for re-identifying vehicles. Together, the semantic segmentation masks and vehicle feature embeddings are modeled. By precisely defining the vehicle region, the co-segmentation technique aids in handling occlusions and backdrop clutter. For vehicle matching, similarity scores are calculated using feature embeddings and segmentation masks. Under difficult circumstances with occlusions and background clutter, VOC-ReID has shown effectiveness in improving matching accuracy.

The overall framework of the proposed VOC-ReID system



Although deep Convolutional Neural Networks (CNN) have demonstrated impressive performance in vehicle re-ID over time [6,7,8], several problems still exist. The viewpoint variation problem in vehicle re-ID is the topic of the fourth technique, Orientation-Aware Vehicle Re-Identification. By adding orientation information during feature extraction, the suggested technique learns orientation-aware features. Each branch of the model's multi-branch architecture oversees extracting features from a particular vehicle orientation. The model captures more robust and discriminative representations for matching cars with various views by taking orientation information into account. Experimental assessments on test datasets show the superiority of the orientation-aware strategy, delivering considerable performance gains over the conventional approaches.

Architecture of proposed framework



Convolutional neural networks (CNNs) are used to extract features using the Siamese neural network method. It excels at tolerating changes in viewpoint and recording local visual patterns.

Transformer-based models, on the other hand, can handle complicated variations and retain semantic links between various components of the vehicle because they are able to collect global contextual information and long-range dependencies. Transformer-based models have shown to perform better in situations with severe occlusions and viewpoint changes.

The question and gallery photos are processed individually in two streams by the Two-Stream Siamese Neural Network architecture. Each stream extracts visual features on its own, enabling better variance handling and discriminating. Contrarily, single-stream designs use a unified feature representation, like transformer-based models. Two-stream architectures are better suited for scenarios with considerable viewpoint changes and occlusions. The choice of architecture depends on the specific needs and complexity of the vehicle re-identification task.

Co-segmentation, which jointly models' segmentation masks and feature embeddings, is a concept introduced by VOC-ReID. Even with occlusions and backdrop clutter, co-segmentation aids in precisely localizing the vehicle region. By concentrating on discriminative regions and reducing irrelevant information, this method increases matching accuracy. Other methods, however, concentrate exclusively on feature extraction from vehicle photos. Co-segmentation or feature extraction should be chosen depending on the dataset's characteristics and the difficulties associated with vehicle re-identification tasks.

All four techniques show increased performance in vehicle re-identification as compared to earlier approaches in terms of efficacy. On benchmark datasets, they achieve cutting-edge or competitive outcomes, demonstrating the viability of their suggested strategies. Remember that the dataset, measurement standards, and experimental design can all affect how effective a method is. Here are a few benefits and drawbacks.

Advantages:

- Its capacity to include both static and dynamic motion inputs, which can increase the model's capacity for discrimination. Additionally, the two-stream design improves the accuracy of the re-identification process by allowing for a more thorough depiction of vehicle photos.
- Transformers' ability to record distant dependencies for TransReID can be useful for simulating complex relationships in vehicle graphics. This may result in enhanced robustness and performance for re-identification jobs.
- By using co-segmentation data, the VOC-ReID can deal with difficult situations like occlusions and complicated backgrounds. This may result in more reliable and accurate re-identification outcomes.
- The capacity of the Orientation-Aware Vehicle Re-Identification to manage various orientations and views enhances the procedure's accuracy. The model's performance is further improved with the addition of an attention-based technique that enables the model to concentrate on discriminative areas in vehicle images.

Disadvantages:

- The limited investigation of more complex approaches, such as multi-scale features or attention mechanisms. Limited datasets and relatively smaller-scale tests and evaluations may hinder generalizability.
- The lack of thorough evaluation and comparison with other cutting-edge techniques in the vehicle re-identification field. Limited investigation of architectural modifications or vehicle-specific enhancements.
- Due to the co-segmentation component, computational complexity is relatively higher. Only a limited number of real-time applications and larger datasets have been explored.
- Limited comparison with cutting-edge techniques and evaluation on large-scale datasets. Possible difficulties in generalizing to many real-world settings and variations.

3. Applications in current research:

Due to its many applications, vehicle re-identification has attracted a lot of attention in contemporary research. Vehicle re-identification is widely used in a variety of industries for the following purposes:

Smart Transportation Systems: Re-identification of vehicles is essential for intelligent parking, traffic monitoring, and other smart transportation systems. These technologies can streamline traffic flow, boost security, and boost overall transportation effectiveness by precisely tracking and recognizing cars.

Law Enforcement and Surveillance: Re-identification of automobiles is helpful for law enforcement and surveillance agencies in tracking down and identifying vehicles used in criminal activity. It makes it possible to recognize suspicious vehicles on several security cameras, assisting with investigations and boosting public safety.

Retail and Marketing Analytics: Re-identification of vehicles can offer insightful data for retail and marketing analytics. Businesses can track consumer activity, gauge foot traffic, examine client demographics, and adjust marketing techniques by detecting cars.

Automated Tolling and Access Control: Automated tolling systems use vehicle re-identification so that vehicles can be precisely identified for quick and easy toll collecting. Additionally, it is used in access control systems, enabling approved cars to enter places that are off-limits.

Smart Cities and Infrastructure: The advancement of smart cities is facilitated by vehicle re-identification, which makes smart parking, traffic signal optimization, and resource allocation possible. It encourages the development of smart infrastructures that improve sustainability and urban mobility.

Although the major categories in which vehicle re-identification is used are transportation, security, and related ones, there are also new uses for this technology in other industries. Let's look at how the fields of finance and healthcare are researching and using vehicle re-identification:

Healthcare:

Ambulance Dispatch and Routing: Identifying and tracking ambulance vehicles with the use of vehicle re-identification can lead to better ambulance dispatch and routing during crises. It makes it possible for quicker reaction times and more efficient resource management.

Finance:

Asset Tracking: Tracking and monitoring the movement of high-value assets, such as automobiles used in banking or logistics operations, can be done through vehicle re-identification. Real-time visibility is provided, allowing for efficient asset management and security.

These uses highlight the adaptability and potential influence of vehicle re-identification across a range of industries. With the development of deep learning models and algorithms, we may anticipate more innovation and use of this technology across a variety of sectors, resulting in improved productivity, security, and all-around performance.

4. Potential future developments:

There are several potential future advancements in the subject of vehicle re-identification that could improve the precision, effectiveness, and usefulness of the technology. These prospective improvements could lead to:

- **Advanced Deep Learning Models:** There is potential in the creation of more complex deep learning models made exclusively for vehicle re-identification. To enhance feature representation and matching accuracy, this includes investigating novel designs like attention mechanisms, graph convolutional networks, or generative adversarial networks.
- **Multi-Modal Fusion:** A richer representation of cars can be provided by integrating different modalities, such as visual data from pictures or videos, sensor data, and textual information, which can also improve re-identification performance. Future studies can concentrate on efficient fusion methods and learning algorithms to make use of the complementary data from many modalities.
- **Robustness to Environmental Factors:** A significant area for improvement is increasing the robustness of vehicle re-identification models to diverse environmental circumstances, including changes in weather, illumination, and viewpoint shifts. These difficulties can be overcome with the aid of methods like adversarial training, data augmentation, and domain adaptability.
- **Privacy-Preserving Techniques:** Privacy issues are a big factor because processing and analyzing sensitive data is a part of vehicle re-identification. The privacy of people and their automobiles can be safeguarded via the development of privacy-preserving techniques, such as secure and encrypted feature extraction or differential privacy measures.

The field of vehicle re-identification can continue to advance by addressing these potential future developments, which will result in better performance, greater applicability, and enhanced capabilities for numerous businesses and domains.

Limitations:

There are still certain issues that need to be resolved in vehicle re-identification despite the advances made utilizing deep learning algorithms. These restrictions consist of:

1. **Variability in Environmental Conditions:** Vehicle re-identification systems may encounter difficulties while adjusting to changes in ambient factors including lighting, weather, and occlusions. Vehicle re-identification algorithms can become inaccurate due to poor lighting, intense rain, or obscured perspectives.
2. **Limited Viewpoints and Camera Coverage:** Systems for re-identifying vehicles rely largely on camera coverage and having access to several viewpoints so that vehicles can be photographed from various angles. The efficiency of these systems may be constrained by limited camera coverage or restricted perspectives, particularly in densely populated areas or by extensive monitoring networks.
3. **Privacy Concerns:** Vehicle re-identification systems acquire and process personal data, which raises questions about privacy. It's vital to strike a balance between the necessity for precise identification and privacy protection. The creation of procedures that protect people's privacy and security while retaining the efficiency of the re-identification process is crucial.
4. **Data Imbalance:** The effectiveness of vehicle re-identification models can be impacted by imbalanced datasets, where certain vehicle classes have noticeably less samples than others. For underrepresented vehicle types, unbalanced data can provide biased models and less accurate predictions. This constraint can be addressed using strategies like data augmentation, class balancing, or transfer learning. The procedure of re-identification must be effective.
5. **Computational Resources and Real-Time Constraints:** For training and inference, deep learning models for vehicle re-identification frequently need a lot of computing power. The computational complexity of these models makes real-time implementation difficult. These constraints can be improved by using hardware acceleration techniques and model optimization.

Potential Solutions:

Several alternative strategies can be investigated to solve the deep learning limits of vehicle re-identification, including:

1) **Improved Feature Representation:** Creating sophisticated feature representation methods can improve the models' ability to discriminate. To capture more accurate and reliable representations of cars, methods like attention mechanisms, graph convolutional networks, or self-supervised learning can be investigated.
2) **Robustness to Environmental Variations**: It is essential to create algorithms that are more resistant to environmental changes. Techniques for data augmentation that are

intended to imitate various lighting and weather situations may be used in this. Additionally, domain adaptation techniques that can train the models to adjust to novel environmental conditions can be incorporated to enhance performance.

3) **Multi-Modal Fusion:** The accuracy and robustness of vehicle re-identification can be increased by integrating numerous modalities, including photos, videos, and sensor data (such LiDAR, radar). To efficiently merge data from several modalities, fusion strategies like late fusion, early fusion, or attention-based fusion might be examined.
4) **Privacy-Preserving Methods:** To address privacy issues, privacy-preserving procedures must be developed. To preserve the effectiveness of the re-identification process while retaining the privacy and security of the data, strategies like federated learning, secure multi-party computation, or differential privacy can be investigated.

Citations:

[1] Pirazh Khorramshahi, Neehar Peri, Jun-cheng Chen, and Rama Chellappa. The devil is in the details: Self-supervised attention for vehicle re-identification. In ECCV, pages 369–386. Springer, 2020.

[2] Yifan Sun, Changmao Cheng, Yuhan Zhang, Chi Zhang, Liang Zheng, Zhongdao Wang, and Yichen Wei. Circle loss: A unified perspective of pair similarity optimization. In CVPR, pages 6398–6407, 2020.

[3] Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, and Shengjin Wang. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). In ECCV, pages 480–496, 2018.

[4] Guangcong Wang, Jian-Huang Lai, Wenqi Liang, and Guangrun Wang. Smoothing adversarial domain attack and p-memory reconsolidation for cross-domain person reidentification. In CVPR, pages 10568–10577, 2020.

[5] Guanshuo Wang, Yufeng Yuan, Xiong Chen, Jiwei Li, and Xi Zhou. Learning discriminative features with multiple granularities for person re-identification. In ACMMM, pages 274–282, 2018

[6] Liu, X., Liu, W., Ma, H., Fu, H.: Large-scale vehicle re-identification in urban surveillance videos. In: IEEE International Conference on Multimedia and Expo (ICME). pp. 1–6 (2016)

[7] 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 (ECCV). pp. 869–884. Springer (2016)

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

References:

1. A Two-Stream Siamese Neural Network for Vehicle Re-Identification by Using Non-Overlapping Cameras - https://paperswithcode.com/paper/a-two-stream-siamese-neural-network-for
2. TransReID: Transformer-based Object Re-Identification: https://paperswithcode.com/paper/transreid-transformer-based-object-re
3. VOC-ReID: Vehicle Re-identification based on Vehicle-Orientation-Camera: https://paperswithcode.com/paper/voc-reid-vehicle-re-identification-based-on
4. Orientation-aware Vehicle Re-identification with Semantics-guided Part Attention Network: https://paperswithcode.com/paper/orientation-aware-vehicle-re-identification