Kufa Journal of Engineering
Vol. 17, No. 3, July 2026, P.P. 564-583
Article history: Received 9 August 2025, last revised 11 November 2025, accepted 4 January 2026

PERSON RE-IDENTIFICATION BASED ON DEEP LEARNING NETWORKS: A SURVEY

Zahraa Faisal [1] and Nidhal K. El Abbadi [2]

[1] Computer Science Department, Computer Science and Mathematics College, University of Kufa, Iraq. Email:zahraaf.shouman@uokufa.edu.iq.

[2] Engineering College, University of Cordoba, Iraq, Email:nidhal.elabbad@uokufa.edu.iq.

https://doi.org/10.30572/2018/KJE/170332

ABSTRACT

Person Re-Identification_(ReID) is a crucial task in computer vision with growing importance in security and engineering applications, particularly in surveillance and smart city systems. The hand-crafted feature-based existing approaches that consider texture and color struggle with complex real challenges involving lighting; person pose; and variable backgrounds. This survey offers an updated and focused review of deep learning-based ReID methods, encompassing research from 2020 to 2025. It investigates in-depth the engineering aspects, including system integration, real-time performance, and sensor constraints, which are often overlooked in reviews of earlier work. Techniques discussed in this study involve CNNs and transformers, triplet loss and contrastive learning, GANs, and methods that enhance matching accuracy and generalization. The paper compares recent methods; presenting their strengths and weaknesses, and setting directions for future research. The survey aims to provide a practical reference for engineers and researchers interested in developing robust and scalable ReID systems in real-world environments.

KEYWORDS

Video-based ReID; Surveillance systems; Human identification; Cross-camera tracking; Smart city applications.

1. INTRODUCTION

Person Re-Identification (ReID) is now the backbone of computer vision research due to its critical applications in intelligent surveillance, smart cities, and retail analysis. The fundamental goal of ReID is to correctly recognize the same person in different images or frames of the video, even if recorded under different lighting, camera angles, and scenarios. The task is inherently difficult because it needs to cope with the vast variation of human appearance, occlusions that happen very frequently, and the challenge created by similar-looking individuals. With the rapid expansion of the world's surveillance networks, the demand for scalable and robust ReID systems has expanded exponentially (Ye et al., 2021). The general process of person Re-ID is illustrates by Fig. 1.

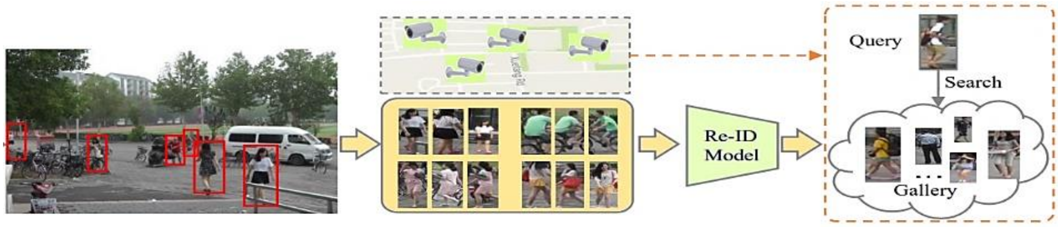


Fig. 1. Illustrate of a person re_identifcation system

Previous Re-ID systems relied primarily on handcrafted features such as color histograms; textures, and spatial layouts. While these approaches laid the basis of the ReID field, the conventional approaches were not effective in dealing with the complexities of the real world, i.e., pose variations, dynamic lighting, and occlusions. Moreover, the traditional approaches lacked good generalization over the datasets, emphasizing their practical shortcomings (Yadav and Vishwakarma, 2020).

Deep learning has transformed ReID with automatic feature extraction and high-level pattern recognition. "Convolutional Neural Networks (CNNs)" have played a central role in the high-level discriminative feature extraction that has resulted in dramatic accuracy and robustness enhancements. Transformer-based models have recently advanced the field further with the successful modeling of the global dependencies of the image and the evasion of the limitations of traditional convolutional modeling. These advances have enabled ReID systems to operate reliably in more challenging and diversified situations (Ming et al., 2022).

Despite these advances, several challenges remain. One of the major issues is domain shift, where models learned on one dataset do not generalize to others. furthermore, real_time application of ReID systems is typically hindered by high computational complexity. Occlusions and intra-class variations also pose significant hurdles, necessitating solutions that can generalize effectively across diverse scenarios while maintaining real_world performance standards(Ning-et al., 2024).

This article reviews recent developments in deep_learning based Re_ID, addresses the challenges hindering their widespread adoption, and explores future trends in this field. By providing detailed and in_depth insights, this study aims to support the development of flexible and scalable Re_ID systems tailored to fit real-world, everyday applications.

The structure of the remainder of this study is as follows:- Section 2 outlines the real-world applications of person re_identifications. Section 3 discusses the main challenges associated with this field. Section 4 reviews of the most significant strategies published in the literature over the past five years. Section 5 describes the commonly used datasets. Section 6 introduces Evaluation metrics that support performance assessment. Section 7 presents the main person Re_ID methods based on deep learning. Section 8 outlines current trends, and future research directions are presented. Finally, Section 9 provides a conclusion.

2. THE REAL WORLD APPLICATIONS OF PERSON RE-IDENTIFICATION.

Section Seven presents the main person Re_ID methods based on deep learning. Section Nine outlines current trends, and future research directions are presented in Section Eight. Finally, Section Nine provides a conclusion.

Track individuals through various camera views in highly secure places such as airports; train stations, or public areas to prevent unauthorized access or criminal activities (Khan et al., 2020). Assists in tracking patients and staff in hospitals to deliver timely medical attention and prevent the risk of patients becoming lost or accessing off-limits areas (Chen, 2018).

Public Safety:-crowd counting and behavior monitoring in public protests or gatherings to prevent overcrowding and enable efficient emergency response. Enhances disaster management by monitoring evacuation procedures and identification of bottlenecks. (Balasubramanian and Nagananthini, 2017).

Urban Planning and Infrastructure Development:_provides insights into the movements of public space of individuals; which aid in planning improved transportation systems, pedestrian paths, and city designs. Helps in optimizing camera locations to provide complete coverage (Carmona, 2019).

Academia and Research: Supports research in AI and computer vision through providing a robust platform on which to test algorithms in realistic surveillance scenarios. Supports data gathering for behavior research and public policymaking (Chen, 2018).

Retail and Commercial Spaces: tracks customer movements and shopping patterns to optimize store layouts and improve customer experiences. Enhances security in malls and shopping centers by identifying potential threats (Chen, 2018).

Human-robot interaction: person Re-IDhas gained increasing attention in robotic vision research. A personal assistant robot relies on person Re-ID algorithms to recognize and give optimal personalized responses to each nearby person (Chen, 2018).

Fig. 2. illustrates the redefinition applications in smart surveillance and people tracking via cameras.

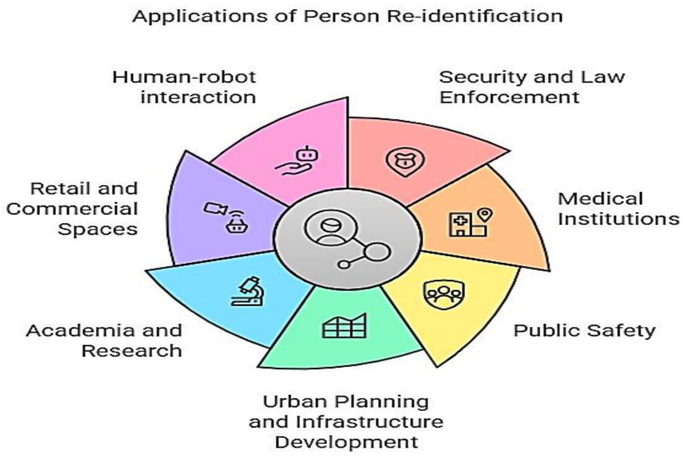


Fig. 2. Applications of a person re-identification.

3. CHALLENGES OF PERSON RE-IDENTIFICATION

Several methods have been developed in recent years to improve the performance of a person's Re-ID. However; most of the challenges, like occlusion, lighting variations, camera view transitions, pose changes, and similar clothes, are yet to be overcome, which further restricts the practical applications of these algorithms in reality.

Addressing these challenges requires the development of large-scale person Re-ID_datasets that include diverse real-world conditions. Due to the evolution from traditional feature extraction to deep learning methods, large-scale datasets have grown considerably, with substantial demands in deep neural network training. However, these large-scale datasets can have huge differences in types and annotation methods, especially in image-based person Re-ID datasets.

Re-identification involves some challenges that hinder its effective performance. These challenges can be considered as a result of variations in appearance, environment, as well as viewpoints, including occlusion, pose, cluttered background, scale difference, illumination variability, viewpoint variability, and low resolution.(Zahra et al.; 2023).

These challenges highlight the need for advanced algorithms, which can cope well with variability, generalize well over a whole set of situations, and also function well in real-world applications. Fig. 3. shows samples of real-world challenges; the Market-1501 dataset was the source of all the samples (Zheng et al., 2015).

Fig. 3. Dataset samples illustrating key challenges in person re-identification.

4. LITERATURE REVIEW

A number of deep learning techniques have been introduced in person Re-ID, focusing on the main challenges of occlusion, scale changes, background clutter, as well as pose variations. This section provides a discussion on the person Re-ID techniques introduced from 2020 to 2025. Each approach is examined based on its main architecture, the main techniques it has utilized, such as multi-branch networks, attention mechanisms, feature fusion techniques, as well as the contribution it has made to the challenge of person Re-ID. Moreover, a comparison of these techniques is provided in Table 1., highlighting their strengths, limitations, used datasets, and performance evaluation metrics. A number of deep learning techniques have been introduced in person Re-ID, focusing on the main challenges of occlusion, scale changes, background clutter, as well as pose variations. This section provides a discussion on the person Re-ID techniques introduced from 2020 to 2025. Each approach is examined based on its main architecture, the main techniques it has utilized, such as multi-branch networks, attention mechanisms, feature fusion techniques, as well as the contribution it has made to the challenge of person Re-ID. Moreover, a comparison of these techniques is provided in Table 1. This systematic comparison enables the identification of the evolution of Re-ID research and points out further directions for improvement.

(Chen et al., 2020) Proposed the "Spatial and Channel Partition Representation Network (SCR)" to improve person re-identification by integrating both global/ local feature extraction. Built on a modified ResNet-50, this network divides feature maps into spatial and channel partitions to capture both coarse and fine-grained details. It uses global max-pooling to extract the most distinctive features. During training, softmax cross-entropy loss is used for local features, while triplet hard loss is applied to global features, making the model more resilient to changes and occlusion in challenging ReID scenarios.

(Wang et al., 2020) Uses ResNet-50 to combine multi-scale features with a Scale Normalization Block. Part-based-pooling to extract more detailed information is combined with the use of global-average-pooling to represent overall information. Moreover, the Normalized Softmax Layer with exclusivity regularization suppresses the redundancy and improves the discrimination of features. This model has high accuracy and robustness when evaluated on '3' datasets.

(Saber et al., 2021) Suggested a model that collects"VGG_16 for feature extraction"and SVC for classification to improve person Re-ID. Feature extraction is performed from the first fully connected layer of VGG_16, and SVC classifies with high durability and accuracy. The method was tested on three datasets and achieved a "Rank1 accuracy" of 89.5% onMarket_1501, demonstrating its effectiveness and efficiency for Re_ID applications.

(Yang et al., 2022) Proposed a model for enhancing person Re-ID that is based on a two-branch CNN with ResNet-50 as the backbone. The 'global_branch' captures holistic features, while the 'local_branch' concentrates on fine-grained details. The salient improvements include global maximal pooling for dimension reduction and triplet loss combined with softmax loss for accurate classification. This model has high accuracy and robustness when evaluated on two datasets.

(Gharbi et al., 2023) Proposed integrates characteristics of CNN for deep feature learning, Local Maximal Occurrence (LOMO) for handling lighting and viewpoint changes, and Gaussian of Gaussian (GOG) for extracting fine details. The features are represented as third-order tensors to enable the fusion of spatial and semantic information for making the representation more powerful. Tensors such as "CNN+LOMO" and "CNN+GOG" are combined in pairs to enhance performance even more. "Tensor-based Cross-View Quadratic Discriminant Analysis (TXQDA)" is also applied to project the tensors to a lower discriminative subspace for effectively distinguishing between the identities.

(Zhu et al., 2023):proposed NDMF-Net that integrates global feature extraction and attention mechanisms with spatial dependency exploration. Using the ResNet-50 backbone, the network applies "Part-Level Hybrid Attention Modules ('PHAM')" to eliminate noise and highlight important regions and "Neighboring-Part Dependency Exploration Modules ('NDEM')" to learn dependency between nearby parts. These features are then fused to generate the final discriminative representation that is more accurate and stable. The method achieves promising performance on three datasets and indicates its effectiveness in addressing challenges such as background noise and pose variations.

(Hussen and Abed, 2023)-The paper proposes a novel method of re-identification of videos via the application of YOLOv5 and OSNet. The model is evaluated on the MARS_dataset, and the results are promising, verifying the efficacy of the application of YOLOv5 as a fast and efficient solution compared to conventional models.

(Wang et al., 2023)-A "Multi-Branch Feature Fusion Network (MFF_Net)" is proposed for person re-identification, integrating global; hybrid local, and saliency_guided branches to enhance robustness and feature diversity. The global branch uses a Swin Transformer with multi_scale feature fusion, while the hybrid local branch addresses illumination and scale variations through attention mechanisms. A saliency-guided branch, reinforced by adversarial classifiers and occlusion simulation, improves occlusion handling. Multi-loss functions, including cross-entropy, triplet loss, and KL divergence, ensure effective training and inter-branch knowledge sharing.

(Zhu et al., 2024) Nevertheless, existing methods predominantly pay attention to the salient parts within the whole image, neglecting some intrinsic occluded features inherent to pedestrians themselves. Driven by this concern, we put forward a new person ReIDnetwork. Our method incorporates the edge features of pedestrians into the representation and makes use of edge information to enable global context feature learning. Moreover, through learning the inner relationships between various parts of pedestrians; we boost the capacity of the network for capturing and learning the inner dependencies within pedestrians and enhance the semantic coherence of the features of pedestrians. Finally, through integrating these multi-aspect features, we create rich and extremely discriminative features of pedestrians, and boost the performance on person Re-ID heavily. As can be seen from experiments, our approach surpasses most state-of-the-arts on person ReID.

(Zhu and Zhang, 2025) proposed a data augmentation framework driven by feature erasing, maintaining fine-grained identity cues at the feature level to enhance person ReID. It applies diagonal swapping augmentation and feature-level erasing for focused suppression of dominant activations while keeping semantic consistency. This technique achieved the best Rank_1 accuracies on three-datasets, proving its robustness and generalization in challenging conditions.

(Wang and Yang, 2025) proposed a "Cross-Attention-Guided Local Feature Enhanced Multi-Branch Network" for person ReID, capturing global and local discriminative features. The three-branch network, including global, partial, and attention-channel, obtains identity-specific cues from multiple perspectives. The achieved results on the Rank-1 accuracy outperform state-of-the-art methods for Market-1501; CUHK03; and DukeMTMC-reID datasets.

Table 1 provides a comparative summary of the performance of person Re-ID methods over the last five years.

New deep learning technologies are driving substantial new advancements in person Re-identification as demonstrated in recent literature. These techniques focus on extracting global/local features, incorporating attention mechanisms and multi-branch networks. We find that the accuracy and robustness while using them on various benchmark datasets have improved considerably. Still, there are problems in applying these methods in real-world situations. Problems like 'background clutter, occlusions, and variations in lighting or pose' continue to be challenges. The variance in results across different datasets indicates models' inability to generalise, which identifies another area of improvement for future research and development. In general, the studies being evaluated are enhancing the understanding of the person re-ID problem and proposing novel methods that are likely to establish a solid foundation for further advances in person re-ID systems. Table 1 features a comparison analysis that offers an insight into strengths and weaknesses of various approaches based on accuracy, feature learning, and performance variations across datasets.

Table 1 presents a comparative analysis of recent person re-identification methods(2020-2025).

Ref.	Pros.	Cons.	Datasets	Rank1%	mAP%
Chen et al, 2020	• Enhance person Re-ID by combining spatial and channel partitions with global features. • Maintains useful secondary features often lost in attention-based models. •Channel partitioning improves robustness by enabling concept-to-concept matching.	• Model complexity increases due to using 13 classifiers, which also slows inference. •Optimal performance needs careful tuning of the loss function's balance parameter (λ).	CUHK03, Market-1501, DukeMTMC	82.2 95.7 91.1	77.6 89.0 81.4
Wang et al., 2020	• By using multi-scale and multi-patch representations, the model captures rich features. • Robust to occlusion and misalignment.	•More computational overhead. • Sensitive to patch and scale selection. • Using limited datasets may lead to overfitting.	CUHK03, Market-1501, DukeMTMC	70.1 93.7 84.4	67.2 81.2 70.4
Saber et al., 2021	• Overcome the data limitation by applying deep transfer from large auxiliary datasets. • Prevent overfitting and improve transfer effectiveness by applying a two-step fine-tuning process. • Provides a new unsupervised co-training method that effectively trains without labels on target data.	• Because the model depends on auxiliary datasets, performance is notably reduced when the target domain highly differs. • Unsupervised co-training may be unstable when data distributions vary widely, because it relies on the quality of pseudo-labels.	CUHK01, Market-1501, i-LIDS	86.1 89.5 75.9	- - -

Ref.	Pros.	Cons.	Datasets	Rank1%	mAP%
Yang et al., 2022	• A dual-branch CNN extracts global and local features for stronger discriminative representation. •Learning efficiency is enhanced when using an adaptive triplet loss over classic triplet loss.	• Using the dual-branch architecture increases the computation cost. • Additional hyperparameters are introduced due to using the adaptive triplet loss, making optimization and tuning more challenging.	CUHK03, Market-1501	90.6 93.8	84.6 -
Gharbi et al., 2023	•Discriminative power improved due to combining high-level CNN features with LOMO and GOG descriptors. •Similarity assessments enhanced by applying Mahalanobis distance for matching.	• The possibility of overfitting and long training times increased due to the use of tensor-based approaches, which tend to have high parameter counts. • Dependency on predefined descriptors (LOMO, GOG) may limit adaptability when compared to end-to-end learning approaches.	VIPeR, PRID450s	49.72 68.09	- -
Zhu et al., 2023	• This method emphasizes fine-grained local features through part-level hybrid attention. • Integrates local and global features into a unified, robust representation.	• Horizontal splitting into parts may suffer when misalignment or variable pose occurs (not explicitly addressed). • Computational cost increases due to the architectural complexity added by "Part-Level Hybrid Attention Modules and Neighboring-Part Dependency Exploration Modules".	DukeMTMC-ReID, Market-1501, MSMT17	89.7 95.8 78.4	79.8 88.9 54.5
Hussen and Abed 2023	• Extract features for real-time person Re-ID by combining YOLOv5 detection with OSNet omni.	• The suggested model using masked faces and temperature measurement may not generalize to standard Re-ID setups.			
Wang et al., 2023	• Using a multi-branch architecture makes it effectively address the challenge of scale, illumination, and occlusion.	• Lacks handling for real-world challenges, limited to common complex scenarios. • Compared to some current methods, the suggested method slightly underperforms in simpler scenes.	Market-1501 DukeMTMC-ReID,	95.2 92.5	85.8 80.6
Zhu et al., 2024	• Using edge information improves feature representation. • Model enhances handling with occlusion and pose changes. • Global and local features are fused by applying a dual-branch network.	• Low resolution and noisy images may degrade the edge quality, which reduces the performance.	DukeMTMC-ReID, MSMT17	90.3 79.2	81.4 55.3
Zhu and Zhang, 2025	• Feature scanning enhances identity learning by masking dominant regions. • It increases data diversity and generalization through diagonal patch swapping.	• Diagonal swapping may cause distortions in asymmetric or inconsistent images. • It relies primarily on augmentation without architectural enhancements.	Market-1501, DukeMTMC-ReID, MSMT17	97.8 91.5 82.6	89.6 78.2 58

Ref.	Pros.	Cons.	Datasets	Rank1%	mAP%
Wang and Yang, 2025	• It effectively enhances discriminatory features by amplifying features and integrating parts. • It combines global and local features to build a robust identity representation.	• Cross-attention fusion can be complex and requires fine-tuning when parts are not semantically consistent. •The computational load increases due to independent branches and attention units.	Market-1501, CUHK03, DukeMTMC-reID	96 82.6 90.8	90 79.7 81.9

5. DATASETS

Person re-identification approaches have experienced unprecedented growth in the recent past, driven by the shift from the traditional approaches to feature extraction to deep learning approaches. This shift has increased the need for large databases, the core purpose of which is training and optimizing the performance of deep neural networks. These data collections are characterized by their diversity of data types and characterization approaches, reflecting the task of models that account for the variability of the target person's appearance across cameras. Important to generate rich datasets with diverse real-world conditions, as this aids in enhancing the accuracy of models and bringing their performance closer to real-world situations.

Fig. 4. illustrates the image-based datasets, whereas Fig. 5. presents the video-based datasets. A sample of images from the datasets is shown in Fig. 6. Additionally, Fig. 7. summarizes the highest accuracy achieved by various methods on benchmark video datasets.

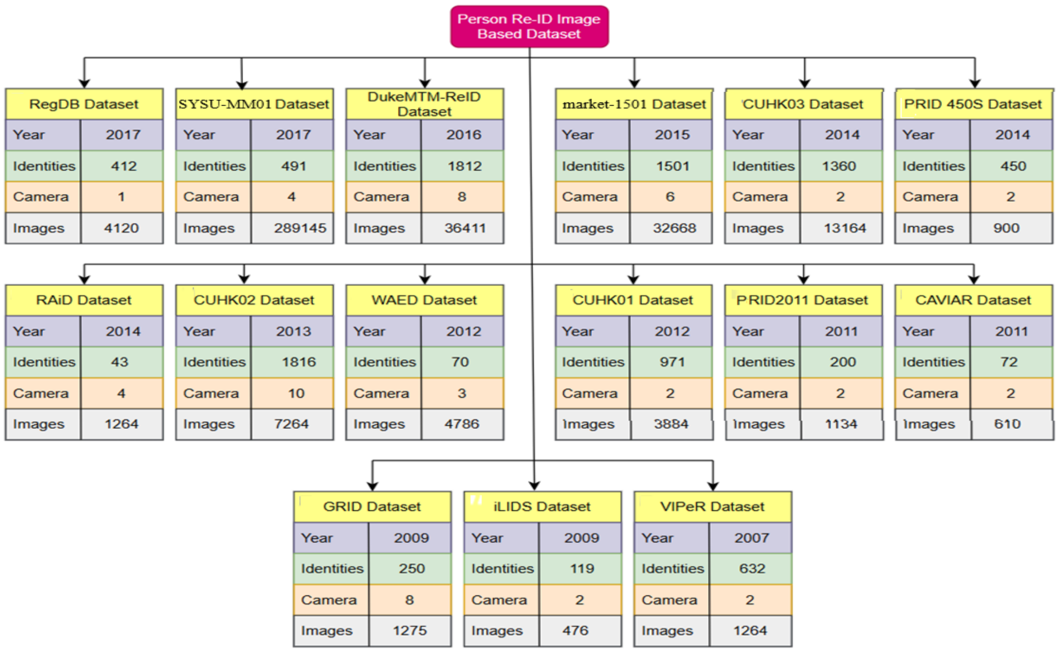


Fig. 4. Overview of commonly used image-based person ReID datasets. (adapted from: Zheng et al., (2015); Zheng et al., (2017); Li et al., (2014); Hirzer et al., (2011); Das et al., (2014); Li and Wang, (2013); Cheng et al., (2011); Loy et al., (2010); Wang et al., (2014); Gray and Tao, (2008).

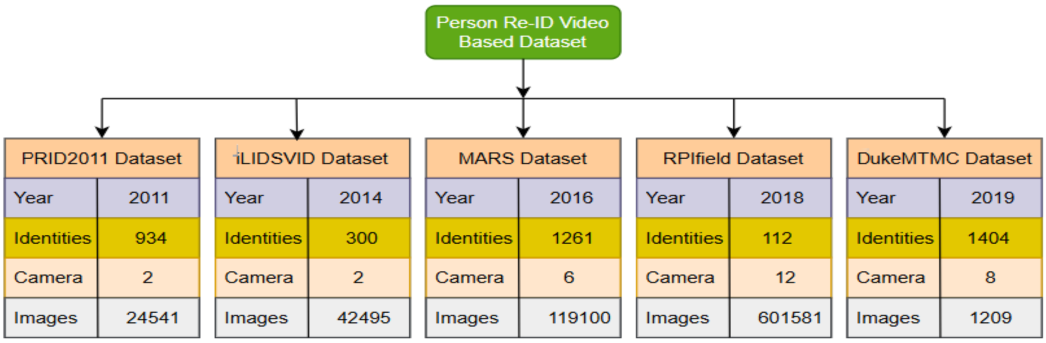


Fig. 5. Overview of commonly used video-based person ReID datasets (adapted from: Hirzer et al. (2011), Wang et al. (2014), and Zheng et al. (2016).

MSMT17_ datasets **DukeMTMC_reID_ datasets** **MARS_ datasets**

Fig. 6. Sampled of person Re-ID datasets.

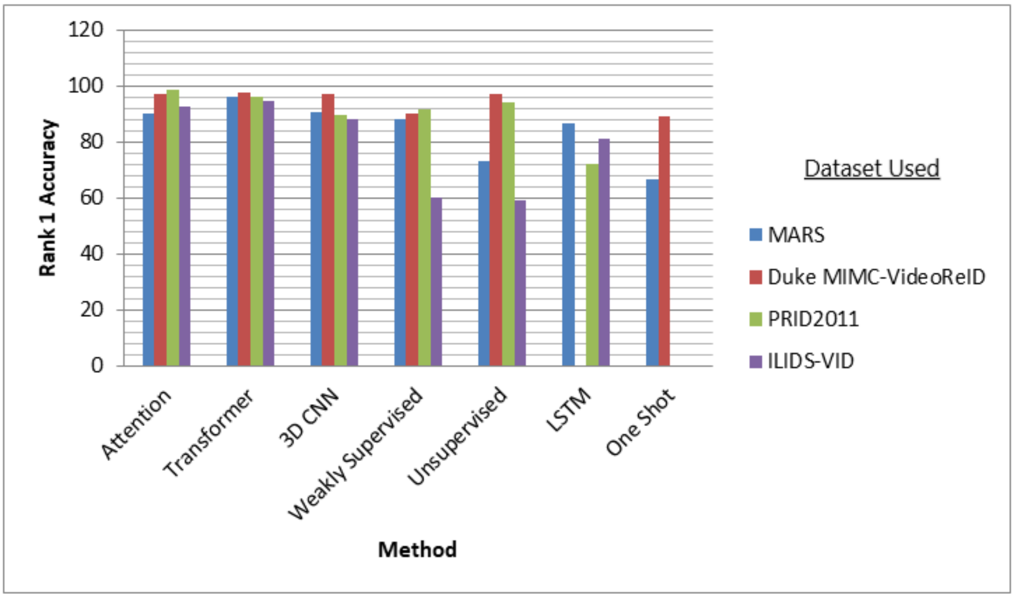


Fig. 7. Highest Achieved Rank-1 Accuracy (%) on Benchmark video Datasets by State-of-the-Art Methods. Based on Saad et al. (2024).

6. EVALUATION MATRICES

This section will introduce the common evaluation metrics utilized to examine the accuracy and performance of pedestrian RE-ID methods in related studies.

6.1. Rank-n Accuracy:

The rank-1 matching rate (rank-1) is a widely used metric that measures the likelihood that the queried image and the image ranked first in similarity within the search dataset correspond to the same target (Bolle et al., 2005) Mathematically, it can be expressed as:

$$\text{Rank -1} = \frac{\sum_{i=1}^{m} s_i}{m} \quad (1)$$

Here, m denotes the total number of images, while s_i is a binary flag variable. If the queried image and the top-ranked image match the same target, $s_i = 1$; otherwise, $s_i = 0$. Generally, a higher rank-1 value indicates better model performance.

Rank-1 is the most important and widely used performance evaluation index since it directly reflects the accuracy of the model.

The (rank-n) indicates the chance that the image being queried is the same as one of the top n most similar images in the search dataset. Common values for n include (1, 3, 5, and 10). The first n results: We check whether there is a correct match amongst the first n results. By contrast, rank-n represents a more comprehensive measure of the model's retrieval ability than rank-1

6.2. Mean Average Precision (mAP):-

Is a metric that has been used to evaluate the performance of models on "ranking and the retrieval" of results, generally for object detection and image retrieval tasks.-It is a measure that combines both 'precision and recall' for the evaluation of relevance and ranking of the retrieved items (Chen et al., 2016).

$$\text{precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{precision} = \frac{TP}{TP + FN} \quad (3)$$

Where:-"true positive (TP), false positive (FP), true negative (TN), and false negative (FN)".

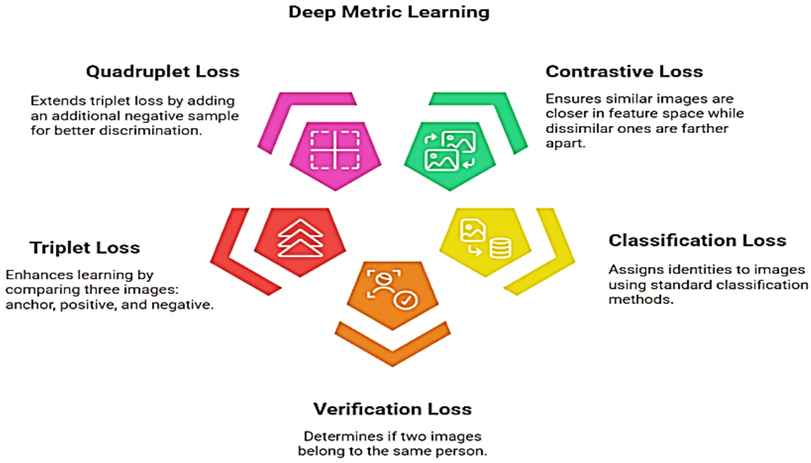
Based on these metrics; a Precision-Recall(P-R) curve is plotted with precision on the verticall axis and recal on the horizontal axis. The areathis curve is referred to as average precision(AP). When AP is averaged across all classes, it results in the mean average precision(mAP); which is commonly used to evaluate overall model performance.

7. RE_ID METHODS BASED ON DEEP LEARNING

This section presents the commonly used evaluation metrics employed to assess the accuracy and overall performance of pedestrian RE-ID methods in previous studies.

7.1. Deep Metric Learning

This technique is designed to learn distance-based correlations between person images by employing various loss functions to quantify image similarity, ensuring that images of the same individual are mapped closer together within the feature space. Fig. 8. provides a comprehensive summary of the most widely adopted loss functions in the literature, as reviewed by (Song et al.(2016), Duan et al.(2017), Ye et al. (2021), Zheng et al. (2019); Hashim and



Mazinani (2025), Zhong et al.(2019), Chen et al.(2017), and (Yassin et al., (2025)).

Fig. 8. Deep Metric Learning Techniques.

7.2. Local Feature Learning

The technique prioritizes local feature extraction over global feature extraction in photos to effectively manage occlusions, stance variations, and background clutter. Fig. 9. shows a summary of the local feature learning techniques (Fu et al. (2019), Miao et al. (2019), Zhang et al. (2020), and Ning et al. (2021)).

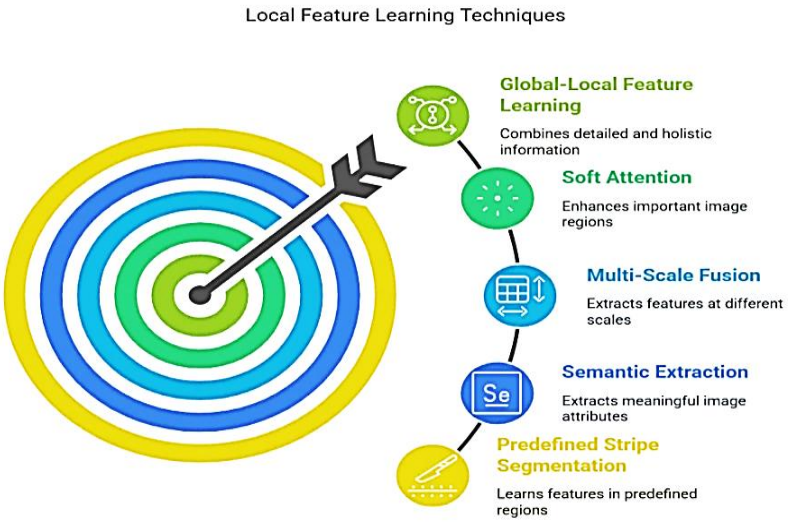


Fig. 9. The local feature learning techniques.

7.3. Generative Adversarial Learning

GANs enhance data by creating realistic variations of training photos, enhancing models' generalization ability to various situations. Fig. 10. shows a summary of GAN techniques (Qian et al. (2018), Zhan and Zhang (2021), and Zheng et al. (2019)).

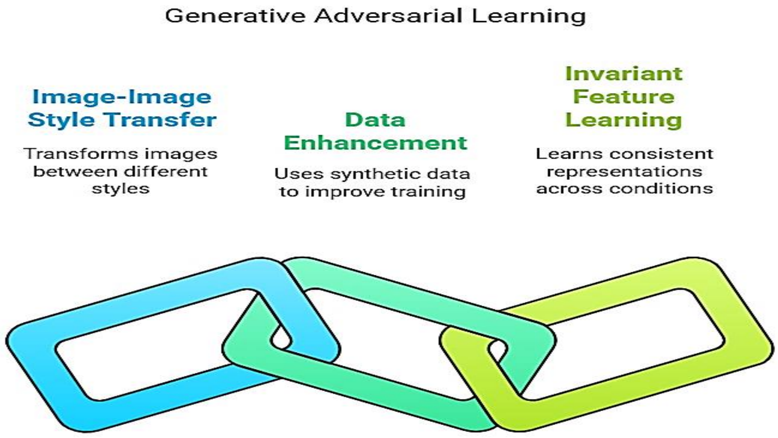


Fig. 10. GANs Learning techniques

7.4. Sequence Feature Learning

The system is designed for human re-identification using video, recognizing a person from multiple frames rather than a single image. Fig. 11. shows a summary of sequence Feature techniques(Chung et al. (2017), Chen et al. (2018), Liao et al. (2018), Hou et al. (2019), Yang et al. (2020), and Hou et al. (2021), (Abbosh et al., (2025)).

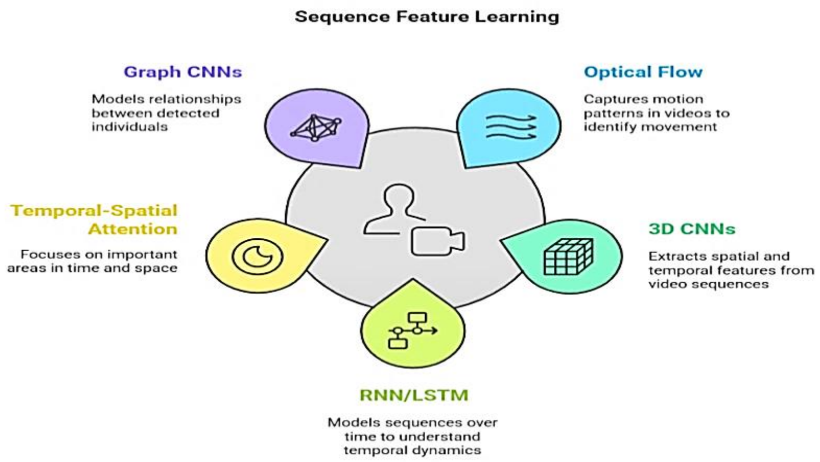


Fig. 11. Sequence Feature learning techniques.

8. TRENDS AND FUTURE RESEARCH DIRECTIONS

Using realistic local data that reflect real-world challenges.

Background removal using advanced models like U-Net to isolate people with minimal background noise,thereby reducing irrelevant information and increasing processing speed during inference.

High-quality training by incorporating features extracted from robust models like ResNet101 and EfficientNetB3.

Using wavelet scattering to extract low-complexity, precise resources.

Achieving high execution speed (faster inference) to make the model suitable for real-time applications.

9. CONCLUSIONS

This paper reviews person Re-ID techniques from 2020 to 2025, which are one of the most important topics in the smart security systems, video surveillance and human tracking. While there has been much improvement in Re-ID, it continues to face concerns like pose differentials, brightness variations, and backgrounds dissimilarity in camera viewpoint.

Deep learning advances have recently gained significant importance for the improved accuracy and representation power of features. In particular convolutional neural networks hybrid attention mechanisms and spatial-semantic modeling. Besides, fresh machine learning creation expedites the Re-ID pipelines while multiplying their real-world applicability. This paper describes the prominent developments, their merits and limitations, and highlights research directions worth pursuing: unsupervised learning, few-shot training, lightweight models for real-time deployment. The results offer a helpful foundation that future endeavors may use to create stronger, more efficient, and more scalable person Re-ID systems.

10. REFERENCES

Abbosh, O.M., Ali, S.M., Ali, D.M. and Alhummada, I.A. (2025) 'Keratoconus detection using deep learning', Kufa Journal of Engineering, 16(2), pp. 280-294. https://doi.org/10.30572/2018/KJE/160217.

Balasubramanian, Y. and Nagananthini, C. (2017) 'Computer vision-based crowd disaster avoidance system: A survey', International Journal of Disaster Risk Reduction, 22, pp. 302-309. https://doi.org/10.1016/j.ijdrr.2017.03.010.

Bolle, R.M., Connell, J.H., Pankanti, S. and Ratha, N.K. (2005) 'The relation between the ROC curve and the CMC', Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05), pp. 15-20. https://doi.org/10.1109/AUTOID.2005.15.

Carmona, M. (2019) 'Principles for public space design, planning to do better', Urban Design International, 24, pp. 47-59. https://doi.org/10.1057/s41289-019-00096-3.

Chen, D., Yuan, Z., Chen, B., Zheng, N. and Yu, K. (2018) 'Video person re-identification with competitive snippet-similarity aggregation and co-attentive snippet embedding', Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1169-1178. https://doi.org/10.1109/CVPR.2018.00127.

Cheng, D.S., Cristani, M., Stoppa, M., Bazzani, L. and Murino, V. (2011) 'Custom pictorial structures for re-identification', Proceedings of the British Machine Vision Conference (BMVC), 1(6), pp. 1-11.

Chen, H., Lagadec, B. and Bremond, F. (2020) 'Learning discriminative and generalizable representations by spatial-channel partition for person re-identification', Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 1-10. https://doi.org/10.1109/WACV45572.2020.9093359.

Chen, W., Chen, X., Zhang, J. and Huang, K. (2017) 'Beyond triplet loss: A deep quadruplet network for person re-identification', Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 403-412. https://doi.org/10.1109/CVPR.2017.50.

Chen, Y. (2018) 'Person re-identification in images with deep learning'. Doctoral dissertation. Université de Lyon. https://tel.archives-ouvertes.fr/tel-02103865.

Chen, Y.C., Zheng, W.S., Lai, J.H. and Yuen, P.C. (2016) 'An asymmetric distance model for cross-view feature mapping in person re-identification', IEEE Transactions on Circuits and Systems for Video Technology, 27(8), pp. 1661-1675. https://doi.org/10.1109/TCSVT.2016.2515991.

Chung, D., Tahboub, K. and Delp, E.J. (2017) 'A two-stream Siamese convolutional neural network for person re-identification', Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1983-1991. https://doi.org/10.1109/ICCV.2017.216.

Das, A., Chakraborty, A. and Roy-Chowdhury, A.K. (2014) 'Consistent re-identification in a camera network', In: Computer Vision - ECCV 2014, Springer, pp. 330-345. https://doi.org/10.1007/978-3-319-10605-2_22.

Duan, Y., Zheng, W.S., Lin, X. and Lai, J.H. (2017) 'Deep localized metric learning', IEEE Transactions on Circuits and Systems for Video Technology, 28(10), pp. 2644-2656. https://doi.org/10.1109/TCSVT.2017.2650895.

Fu, Y., Wei, Y., Huang, G., Shi, H., Wang, X. and Zhang, T. (2019) 'Horizontal pyramid matching for person re-identification', Proceedings of the AAAI Conference on Artificial Intelligence, 33(1), pp. 8295-8302. https://doi.org/10.1609/aaai.v33i01.33018295.

Gharbi, A.A., Harizi, W., Atri, M. and Dey, N. (2023) 'Enhancing Person Re-Identification through Tensor Feature Fusion', arXiv preprint arXiv:2312.10470. available: https://arxiv.org/abs/2312.10470.

Gray, D. and Tao, H. (2008) 'Viewpoint invariant pedestrian recognition with an ensemble of localized features', In: European Conference on Computer Vision (ECCV), Springer, pp. 262-275. https://doi.org/10.1007/978-3-540-88693-8_20.

Hashim, A.A. and Mazinani, M. (2025) 'Detection of keratoconus disease depending on corneal topography using deep learning', Kufa Journal of Engineering, 16(1), pp. 463-478. https://doi.org/10.30572/2018/KJE/160125.

Hirzer, M., Beleznai, C., Roth, P.M. and Bischof, H. (2011) 'Person re-identification by descriptive and discriminative classification', In: Image Analysis: 17th Scandinavian Conference, SCIA 2011, Ystad, Sweden, Proceedings, Springer, pp. 91-102. https://doi.org/10.1007/978-3-642-21227-7_9.

Hou, R., Ma, B., Chang, H. and Shan, S. (2019) 'VRSTC: Occlusion-free video person re-identification', Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7183-7192. https://doi.org/10.1109/CVPR.2019.00736.

Hou, R., Ma, B., Zhang, Y., Chang, H. and Shan, S. (2021) 'Bicnet-TKS: Learning efficient spatial-temporal representation for video person re-identification', Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4288-4297. https://doi.org/10.1109/CVPR46437.2021.00426.

Hussien, S.A.A. and Abed, A.A. (2023) 'Real-time person re-identification using omni-scale feature learning network and Yolov5: A comparative study', Ingénierie des Systèmes d'Information, 28(3), pp. 685-691. https://doi.org/10.18280/isi.280313.

Khan, K., Khan, A., Madani, S.A., Rodrigues, J.J.P.C. and Almogren, A. (2020) 'Advances and trends in real-time visual crowd analysis', Sensors, 20(18), p. 5073. https://doi.org/10.3390/s20185073.

Liao, X., Yu, L., Zheng, L., Zheng, Y. and Gong, S. (2018) 'Video-based person re-identification via 3D convolutional networks and non-local attention', In: Asian Conference on Computer Vision (ACCV), Springer, pp. 620-636. https://doi.org/10.1007/978-3-030-20893-6_38.

Li, W. and Wang, X. (2013) 'Locally aligned feature transforms across views', Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3594-3601. https://doi.org/10.1109/CVPR.2013.461.

Li, W., Zhao, R., Xiao, T. and Wang, X. (2014a) 'DeepReID: Deep filter pairing neural network for person re-identification', Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 152-159.: https://doi.org/10.1109/CVPR.2014.28.

Loy, C.C., Xiang, T. and Gong, S. (2010) 'Time-delayed correlation analysis for multi-camera activity understanding', International Journal of Computer Vision, 90(1), pp. 106-129. https://doi.org/10.1007/s11263-010-0330-5.

Miao, J., Wu, Y., Liu, P., Ding, Y. and Yang, Y. (2019) 'Pose-guided feature alignment for occluded person re-identification', Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 542-551. https://doi.org/10.1109/ICCV.2019.00063.

Ming, Z., Wang, J., Wang, J., Gong, H., Lin, X. and Han, G. (2022) 'Deep learning-based person re-identification methods: A survey and outlook of recent works', Image and Vision Computing, 119, p. 104394. https://doi.org/10.1016/j.imavis.2021.104394.

Ning, E., Yan, W., Liu, F., Liu, J., Huang, Z. and Li, X. (2024) 'Occluded person re-identification with deep learning: A survey and perspectives', Expert Systems with Applications, 239, p. 122419. https://doi.org/10.1016/j.eswa.2023.122419.

Ning, X., Wang, G., Zhang, Y., Zhang, W. and Wang, X. (2021) 'JWSAA: Joint weak saliency and attention aware for person re-identification', Neurocomputing, 453, pp. 801-811. https://doi.org/10.1016/j.neucom.2021.04.078.

Qian, X., Fu, Y., Xiang, T., Wang, W. and Xu, J. (2018) 'Pose-normalized image generation for person re-identification', Proceedings of the European Conference on Computer Vision (ECCV), pp. 650-667. https://doi.org/10.1007/978-3-030-01237-3_39.

Saad, R.S.M., Moussa, M.M., Abdel-Kader, N.S. and Tolba, M.F. (2024) 'Deep video-based person re-identification (Deep Vid-ReID): comprehensive survey', EURASIP Journal on Advances in Signal Processing, 2024(63) https://doi.org/10.1186/s13634-024-01139-x.

Saber, S., Amin, K. and Hammad, M. (2021) 'An efficient person re-identification method based on deep transfer learning techniques', International Journal of Computers and Information, 8(2), pp. 94-99.

Song, H.O., Xiang, Y., Jegelka, S. and Savarese, S. (2016) 'Deep metric learning via lifted structured feature embedding', Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4004-4012. https://doi.org/10.1109/CVPR.2016.435.

Wang, C., Ma, X., Wang, L., Li, W. and Zhang, W. (2020) 'Multi-scale multi-patch person re-identification with exclusivity regularized softmax', Neurocomputing, 382, pp. 64-70. https://doi.org/10.1016/j.neucom.2019.11.097.

Wang, T., Gong, S., Zhu, X. and Wang, S. (2014) 'Person re-identification by video ranking', In: Computer Vision - ECCV 2014, Part IV, Springer, pp. 688-703. https://doi.org/10.1007/978-3-319-10599-4_45.

Wang, T., Gong, S., Zhu, X. and Wang, S. (2014) 'Person re-identification by video ranking', In: European Conference on Computer Vision (ECCV), Springer, pp. 688-703. https://doi.org/10.1007/978-3-319-10593-2_45.

Wang, X. and Yang, J. (2025) 'Cross-Attention Guided Local Feature Enhanced Multi-Branch Network for Person Re-Identification', Electronics, 14(8), p. 1626. https://doi.org/10.3390/electronics14081626.

Wang, X., Hu, X., Liu, P. and Tang, R. (2023) 'A Person Re-Identification Method Based on Multi-Branch Feature Fusion', Applied Sciences, 13(21), p. 11707. https://doi.org/10.3390/app132111707.

Yadav, A. and Vishwakarma, D.K. (2020) 'Person re-identification using deep learning networks: A systematic review', arXiv preprint arXiv:2012.13318. https://doi.org/10.48550/arXiv.2012.13318.

Yang, B., Lin, Z., Wang, Y. and Zhang, X. (2022) 'A feature extraction method for person re-identification based on a two-branch CNN', Multimedia Tools and Applications, 81(27), pp. 39169-39184. https://doi.org/10.1007/s11042-021-11796-4.

Yang, J., Li, J., Xu, H., Wang, L., Zhang, J. and Huang, K. (2020) 'Spatial-temporal graph convolutional network for video-based person re-identification', Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3289-3299. https://doi.org/10.1109/CVPR42600.2020.00335.

Yassin, R.A., Valizadeh, M. and Abdulaal, A.H. (2025) 'A novel deep learning approaches for multi-class histopathological sub-image classification using prior knowledge', Kufa Journal of Engineering, 16(3), pp. 725-755. https://doi.org/10.30572/2018/KJE/160340.

Ye, M., Chen, J., Zhu, X. and Zhang, Q. (2021) 'Collaborative refining for person re-identification with label noise', IEEE Transactions on Image Processing, 31, pp. 379-391. https://doi.org/10.1109/TIP.2021.3132940.

Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L. and Hoi, S.C.H. (2021) 'Deep learning for person re-identification: A survey and outlook', IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(6), pp. 2872-2893. https://doi.org/10.1109/TPAMI.2020.3003457.

Zahra, A., Ullah, A., Khan, A., Rho, S. and Baik, S.W. (2023) 'Person re-identification: A retrospective on domain-specific open challenges and future trends', Pattern Recognition, 142, p. 109669. https://doi.org/10.1016/j.patcog.2023.109669.

Zhan, F. and Zhang, C. (2021) 'Spatial-aware GAN for unsupervised person re-identification', Proceedings of the 25th International Conference on Pattern Recognition (ICPR), IEEE, pp. 1181-1188. https://doi.org/10.1109/ICPR48806.2021.9412144.

Zhang, Z., Lan, C., Zeng, W., Jin, X. and Chen, Z. (2020) 'Relation-aware global attention for person re-identification', Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3186-3195. https://doi.org/10.1109/CVPR42600.2020.00324.

Zheng, L., Bie, Z., Sun, Y., Wang, J., Su, C., Wang, S. and Tian, Q. (2016) 'MARS: A video benchmark for large-scale person re-identification', In: Computer Vision - ECCV 2016, Part VI, Springer, pp. 868-884. https://doi.org/10.1007/978-3-319-46466-1_52.

Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J. and Tian, Q. (2015) 'Scalable person re-identification: A benchmark', Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1116-1124. https://doi.org/10.1109/ICCV.2015.133.

Zheng, M., Wang, S., Wang, H. and Yang, R. (2019) 'Re-identification with consistent attentive siamese networks', Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5735-5744. https://doi.org/10.1109/CVPR.2019.00589.

Zheng, Z., Yang, X., Yu, Z., Zheng, L., Yang, Y. and Kautz, J. (2019) 'Joint discriminative and generative learning for person re-identification', Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2138-2147. https://doi.org/10.1109/CVPR.2019.00224.

Zheng, Z., Zheng, L. and Yang, Y. (2017) 'Unlabeled samples generated by GAN improve the person re-identification baseline in vitro', Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 3754-3762. https://doi.org/10.1109/ICCV.2017.403.

Zhong, Z., Zheng, L., Luo, Z., Li, S. and Yang, Y. (2019) 'Invariance matters: Exemplar memory for domain adaptive person re-identification', Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 598-607. https://doi.org/10.1109/CVPR.2019.00068.

Zhu, C., Wang, Y., Li, L., Zhou, Y. and Gao, Y. (2023) 'Neighboring-Part Dependency Mining and Feature Fusion Network for Person Re-Identification', IEEE Access, 11, pp. 49760-49777. https://doi.org/10.1109/ACCESS.2023.3289379.

Zhu, C., Zhou, W. and Ma, W. (2024) 'Person Re-Identification Network Based on Edge-Enhanced Feature Extraction and Inter-Part Relationship Modeling', Applied Sciences, 14(18), p. 8244. https://doi.org/10.3390/app14188244.

Zhu, S. and Zhang, H. (2025) 'Improving Person Re-Identification via Feature Erasing-Driven Data Augmentation', Mathematics, 13(16), p. 2580. https://doi.org/10.3390/math13162580.