## **Re-Identification in Urban Scenarios**

Subjects: Computer Science, Artificial Intelligence | Computer Science, Interdisciplinary Applications Contributor: Joao Tavares

Multi-object Re-Identification (ReID), based on a wide range of surveillance cameras, is nowadays a vital aspect in modern cities, to better understand city movement patterns among the different infrastructures, with the primary intention of rapidly mitigate abnormal situations, such as tracking car thieves, wanted persons, or even lost children. Given an image or video of an object-of-interest (query), object identification aims to identify the object from images or video feed taken from different cameras.

person ReID computer vision

deep neural networks

image enhancement

## 1. Introduction

The task of object ReID on image cameras has been studied for several years by the computer vision and pattern recognition communities [1], with the primary goal to ReID a query object among different cameras.

Multi-object ReID, based on a wide range of surveillance cameras, is nowadays a vital aspect in modern cities, to better understand city movement patterns among the different infrastructures <sup>[2]</sup>, with the primary intention of rapidly mitigate abnormal situations, such as tracking car thieves, wanted persons, or even lost children.

This is still a challenging task, since an object's appearance may dramatically change across camera views due to the significant variations in illumination, poses or viewpoints, or even cluttered backgrounds  $^{(2)}$  (Figure 1). According to the state-of-the-art research studies, existing object ReID methodologies can be divided into two main categories: image-based and video-based object ReID.



Figure 1. Some common problems found in object ReID.

The former category focuses on matching a probe image of one object, with an image of the object with the same ID among gallery sets, which is mainly established based on image content analysis and matching. In contrast, the latter category focuses on matching two videos, exploiting different information, such as temporal, and motion-based information. A gallery corresponds to a collection of object images gathered from different perspectives over time. In both approaches, the pairs of objects to be matched are analogous. However, in real scenarios, object ReID needs to be conducted between the image and video. For example, given a picture of a criminal suspect, the police would like to quickly locate and track the suspect from hundreds of city surveillance videos. The ReID under this scenario is called image-to-video person ReID, where a probe image is searched in a gallery of videos acquired from different surveillance cameras.

Although videos contain more information, image-to-video ReID share the same challenges with image-based and video-based objects ReID, namely, similar appearance, low resolution, substantial variation in poses, occlusion, and different viewpoints. In addition, an extra difficulty resides on the match between two different datasets, one static and another dynamic, i.e., image and video, respectively.

Image and video are usually represented using different features. While only visual features can be obtained from a single image, both visual features and spatial-temporal features can be extracted from a video. Recently, CNN has shown potential for learning state-of-the-art image feature embedding <sup>[3][4]</sup> and Recurrent Neural Network (RNN) yields a promising performance in obtaining spatial-temporal features from videos <sup>[5][6]</sup>.

In general, there are two major types of deep learning structures for object ReID; namely, verification models and identification models. Verification models take a pair of data as input and determine whether they belong to the same object or not, by leveraging weak ReID labels that can be regarded as a binary-class classification or similarity task <sup>[7]</sup>. In contrast, identification models aim at feature learning by treating object ReID as a multi-class classification task <sup>[4]</sup>, but lack direct similarity measurement between input pairs. Due to their complementary

advantages and limitations, the two models have been combined to improve the ReID performance <sup>[8]</sup>. However, in the image-to-video object ReID task, a cross-modal system, directly using DNN and the information provided by the target task still cannot perfectly bridge the "media gap", which means that representations of different datasets are inconsistent. Therefore, most of the current works directly rely on weights from pre-trained deep networks as the backbone to obtain initial values for the target model and initiate the pre-trained network structure to facilitate the training of the new deep model. **Figure 2** depicts the common base architecture for person ReID.



Figure 2. Common base architecture for person ReID.

## 2. Deep Learning for Object Re-Identification

A turning point in the history of machine learning and computer vision was reached by researchers of the University of Toronto, who proposed a new image classification approach and achieved excellent results in the ImageNet <sup>[9]</sup> competition <sup>[10]</sup>. The winning proposal, defined as AlexNet, consisted of a CNN composed of a set of stacked deep layers and dense layers that enabled the reduction of the error drastically. However, the first appearances of CNN dated from 1990, when Lecun et al. <sup>[11]</sup> proposed a CNN method addressing the task of hand-written digit recognition to alleviate the work of the postal office.

A CNN multi-layer contains at least one layer to perform convolution operations from image inputs, by using filters of kernels that are translated across and down the input matrix, to generate a feature representation map of the original image input. The characteristics of these filters can be widely different, and each one of them is composed of learnable parameters, updated through a gradient descent optimization scheme. The same layer can employ other filters. For the same part of the image, each filter produces a set of local responses, enabling correlation of specific pixel information with the content of the adjacent pixels. After the proposal by <sup>[10]</sup>, many different network architectures were developed to address such a problem, each one with its inner characteristics, to name a few, the VGG <sup>[12]</sup>, Inception <sup>[13]</sup>, and ResNet <sup>[13]</sup> architectures. Many other network architectures were proposed, but in summary, they all share some building blocks.

In regard to the ReID task, the most used backbone model architecture is ResNet <sup>[14]</sup>, due to its flexibility and ease of reusing and implementation to solve new problems. Most of the ReID tasks explore the use of pre-trained ResNet as backbones for feature extraction for the object ReID task.

ResNets can address the vanishing gradient problem when the networks are too deep, making the gradients quickly shrink to zero after several chain rule applications, leading toward "not updating" the weights and, therefore, harming of the learning process. With ResNets, the gradients can flow directly through the skip connections backward from later layers to the initial filters.

ResNets can have different sizes, depending on how big each model layer is, and how many layers it contains. As an example, ResNet 34 <sup>[15]</sup> contains one convolution and pooling layer step, followed by four similar layers (**Figure 3**). Each layer follows the same pattern by performing

convolutions with a fixed feature map dimension, and by bypassing the input every two convolutions. A concern with the special characteristics of ResNet 34 is the fact that the width *W* and height *H* dimensions remain constant during the entire single layer. The size reduction is achieved by the stride size used in the convolution kernels, instead of the polling layers commonly used in other models.



Figure 3. A common ResNet block architecture.

Every layer of a ResNet is composed of several blocks, enabling it to go deeper. This deepness is achieved by increasing the number of operations within a block, while maintaining the total number of layers. Each operation comprises a convolution step, batch normalization, and a ReLU activation to a particular input; except for the last operation of the block, which does not contain a ReLU activation function.

## **3. Evaluation Metrics**

Cumulative Matching Characteristic curve (CMC) is a common evaluation metric for person or object ReID methods. It can be considered a simple single-gallery-shot setting, where each gallery identity only has one instance. Given a probe image, an algorithm will rank the entire gallery sample according to the distances to the probe, with the CMC top-*k* accuracy given as:

# $Acc_k = egin{cases} 1 & ext{if top} -k ext{ ranked gallery samples contain the query identity,} \ 0 & ext{otherwise,} \end{cases}$

which is a shifted step function. The final Cumulative Matching Characteristics (CMC) curve is built by averaging the shifted step functions over all the queries.

Another commonly used metric is the mean Average Precision (mAP), which is very often employed on each image query, and defined as:

$$MAP = rac{\sum_{q=1}^Q AveP(q)}{Q},$$

where Q is the number of queries.

## 4. Person Re-Identification

Person ReID is the problem of matching the same individuals across multiple image cameras or across time within a single image camera. The computer vision and pattern recognition research communities have paid particular attention to it due to its relevance in many applications, such as video surveillance, human–computer interactions, robotics, and content-based video retrieval. However, despite years of effort, person ReID remains a challenging task for several reasons <sup>[16]</sup>, such as variations in visual appearance and the ambient environment caused by different viewpoints from different cameras.

Significant changes in humans pose—across time and space—background clutter and occlusions; different individuals with similar appearances present difficulties to the ReID tasks. Moreover, with little or no visible image faces due to low image resolution, the exploitation of biometric and soft-biometric features for person ReID is limited. For the person ReID task, databases and different approaches have been proposed by several authors.

#### 4.1. Person Re-Identification Databases

The recognition of human attributes, such as gender and clothing types, has excellent prospects in real applications. However, the development of suitable benchmark datasets for attribute recognition remains lagged. Existing human attribute datasets are collected from various sources or from integrating pedestrian ReID datasets. Such heterogeneous collections pose a significant challenge in developing high-quality fine-grained attribute recognition algorithms.

Among the public databases that have been proposed for person ReID, some examples can be found in the open domain, such as the Richly Annotated Pedestrian (RAP) <sup>[17]</sup>, which contains images gathered from real multicamera surveillance scenarios with long-term collections, where data samples are annotated, not only with finegrained human attributes, but also with environmental and contextual factors. RAP contains a total of 41,585 pedestrian image samples, each with 72 annotated attributes, as well as viewpoints, occlusions, and body part information.

Another example is the VIPeR <sup>[18]</sup> database, which contains 632 identities acquired from 2 cameras, forming a total of 1264 images. All images were manually annotated, with each image having a resolution of 128 x 48pixels.

A summary of the public available databases for person ReID, with main characteristics, is presented in **Table 1**.

**Table 1.** Global overview of the found public available databases for person ReID.

Dataset	# Identities	# Cameras	# Images	Label Method	Size	Tracking Sequences
VIPeR	632	2	1264	Hand	128  imes 48	NO
ETH1,2,3	853,528	1	8580	Hand	Vary	YES
QMUL ILIDS	119	2	476	Hand	Vary	NO
GRID	1025	8	1275	Hand	Vary	NO
CAVIAR4reid	72	2	1220	Hand	Vary	NO
3DPeS	192	8	1011	Hand	Vary	NO
PRID2011	934	2	24,541	Hand	$128\times 64$	YES
WARD	70	3	4786	Hand	128  imes 48	YES
SAIVT-Softbio	152	8	64,472	Hand	Vary	YES
CUHK01	971	2	3884	Hand	160  imes 60	NO
CUHK02	1816	10 (5 pairs)	7264	Hand	160  imes 60	NO
CUHK03	1467	10 (5 pairs)	13,164	Hand	Hand/DPM	NO
RAiD	43	4	6920	Hand	$128\times 64$	NO
iLIDS-VID	300	2	42,495	Hand	Vary	YES
MPR Drone	84	1	-	ACF	Vary	NO
HDA Person Dataset	53	13	2976	Hand/ACF	Vary	YES
Shinpuhkan Dataset	24	16	-	Hand/ACF	128  imes 48	YES
CASIA Gait Database B	124	11	-	Background subtraction	Vary	YES
Market-1501	1501	6	32,217	Hand/DPM	$128\times 64$	NO
PKU-reid	114	2	1824	Hand	$128\times 64$	NO
PRW	932	6	34,304	Hand	Vary	NO
Large scale person search	11,934	-	34,574	Hand	Vary	NO
MARS	1261	6	1,191,003	DPM+GMMCP	256  imes 128	YES
DukeMTMC-reid	1812	8	36,441	Hand	Vary	NO
DukeMTMC4reid	1852	8	346,261	Doppia	Vary	NO
Airport	9651	6	39,902	ACF	$128\times 64$	NO
MSMT17	4101	15	126,441	Faster RCNN	Vary	NO
RPIfield	112	12	1,601,581	ACF	Vary	NO

DPM—deformable part models, ACF—pyramid features, GMMCP—generalized maximum multi clique.

Although many other datasets suitable for object ReID can be found, the ones listed in **Table 1** are widely used by most of the authors as benchmarks for performance evaluation and comparison of the proposed works.

#### 4.2. Person Re-Identification Methods

In this section, deep learning-based person ReID methods are grouped into four main categories, as represented in **Figure 4**, including methods for feature learning, sequence learning, generative learning, and deep learning metrics. These categories encompassed several methods, and they are discussed in the following in terms of their main aspects and experimental results.



Figure 4. Deep learning-based person re-identification methods.

## 5. ReID and Spatial–Temporal Multi Object ReID Methods

One of the main difficulties of object ReID is to operate in distributed scenarios and account for spatial-temporal constraints for multi-object ReID. Main techniques explore the use of RNN models to construct tracklets to assign IDs to objects, enabling to robustly handle occlusions. In contrast, others rely on 3D CNN to attain temporal dependencies of the object in track.

#### 5.1. Multi Object ReID Datasets with Trajectories

Multi-Object ReID considers attributes, such as shape and category combined with trajectories, and is one of the objectives that go towards the objective to perform multi-object ReID in a real urban scenario. However, the development of suitable benchmark datasets for attribute recognition remains sparse. Some object ReID datasets

contain trajectories collected from various sources, and such heterogeneous collection poses a significant challenge in developing high-quality fine-grained multi-object recognition algorithms.

Among the publicly available datasets for ReID, one example is the NGSIM dataset <sup>[19]</sup>, a publicly available data set with hand-coded Ground Truth (GT) that enables evaluating multi-camera, multi-vehicle tracking algorithms on real data, quantitatively. This dataset includes multiple views of a dense traffic scene with stop-and-go driving patterns, numerous partial and complete occlusions, and several intersections.

Another example is the KITTI Vision Benchmark Suite <sup>[20]</sup>, which is composed of several datasets for a wide range of tasks, such as stereo, optical flow, visual odometry, 3D object detection, and 3D tracking, complemented with accurate ground truth provided by Velodyne laser 3D scanner and real GPS localization system, **Figure 5**. The datasets were captured by driving around the mid-size city of Karlsruhe, in Germany, in rural areas, and on highways, and on average, there are up to 15 cars and 30 pedestrians per image. A detailed evaluation metric and evaluation site are also provided.



Figure 5. Example of the multi-object tracking and segmentation system (MOTS).

A summary of the public available databases for multi-object car ReID with trajectories is presented on **Table 2**.

Table 2. Global overview of the public available databases for multi-object car ReID with trajectories.

Dataset	# Identities	# Cameras	# Images	Label Method	Size	Tracking Sequences
NGSIM						
KITTI	-	-	-	R-CNN	1392  imes 512	Yes
UA-DETRAC	825	24	1.21 M	Manual	960 imes 540	Yes
VehicleID	26,267	-	221 K	Manual	Vary	Yes
VeRi-776	776	18	50 k	Manual	Vary	Yes
CompCar	1687	-	18 k	Manual	Vary	Yes
PKU-Vehicle	-	-	18 M	Manual	Cropped	Yes
MOT20-03	735	-	356 k	R-CNN	1173 imes 880	Yes
MOT16	-	-	476 k	R-CNN	$1920\times1080$	Yes
TRANCOS	46,796	-	58 M	HOG	Vary	Yes
WebCamT	-	212	60 k	-	-	No

R-CNN—region proposals with CNN, HOG—histogram oriented gradients.

In addition, the PASCAL VOC project <sup>[21]</sup>, provides standardized image datasets for object class recognition, with annotations that enable evaluation and comparison of different methods. Another useful dataset is ImageNet <sup>[10]</sup>. This image database is organized according to the WordNet hierarchy (currently only the nouns), where each node is represented by hundreds and thousands of images. Currently, each node contains an average of over 500 images. The dataset encompasses a total of 21,841 non-empty sets, forming a total number of images of around 14 million, with several images with bounding box annotations of 1 million. In addition, it contains 1.2 million images with pre-computed SIFT features <sup>[22]</sup>. Pre-trained networks on these two datasets are commonly reused as backbones for feature extraction to perform several tasks, such as object ReID.

#### 5.2. Spatial-Temporal Constrained and Multi Object ReID Methods

In this section, deep learning-based vehicle ReID methods are grouped into four main categories, as represented in **Figure 6**; which includes methods for feature learning, sequence learning, deep learning metrics, and tracking, with these categories encompassing several methods. The main aspects of each category method and their experimental results are discussed in the following.





### 6. Methods for Image Enhancement

Severe weather conditions, such as rain and snow, adversely may affect the visual quality of the images acquired under such conditions; thus, rendering them useless for further usage and sharing. In addition, such degraded images usually drastically affect the performance of vision systems. Mainly, it is essential to address the problem of single image de-raining. However, the inherent ill-posed nature of the situation presents several challenges.

In <sup>[23]</sup>, it is proposed an image de-raining conditional generative adversarial network (ID-CGAN) that account for quantitative, visual, and also discriminative performance into the objective function. The proposal method explores the capabilities of conditional generative adversarial networks (CGAN), in combination with additional constraint to enforce the de-rained image to be indistinguishable from its corresponding GT clean image. A refined loss function and other architectural novelties in the generator–discriminator pair were also introduced, with the loss function aimed towards the reduction of artifacts introduced by GAN, ensuring better visual quality. The generator subnetwork is constructed using densely connected networks, whereas the discriminator is designed to leverage global and local information and between real/fake images. Exhaustive experiments were conducted against several State-of-the-Art (SOTA) methods using synthetic datasets derived from the UCID <sup>[24]</sup> and BSD-500 <sup>[25]</sup> datasets, and with external noise artifacts added. The experiments were evaluated on synthetic and real images using several evaluation metrics such as peak signal to noise ratio (PSNR), structural similarity index (SSIM) <sup>[26]</sup>.

universal quality index (UQI) <sup>[27]</sup>, and visual information fidelity (VIF) <sup>[28]</sup>, with the proposed model achieving an PSNR (DB) of 24.34. Moreover, experimental results evaluated on object detection methods, such as FasterRCNN <sup>[29]</sup>, demonstrated the effectiveness of the proposed method in improving the detection performance on images degraded by rain.

A single-image-based rain removal framework was proposed in <sup>[30]</sup> by properly formulating the rain removal problem as an image decomposition problem based on the morphological decomposition analysis. The alternative to applying a conventional image decomposition technique, the proposed method first decomposes an image into the low and high-frequency (HF) components by employing a bilateral filter. The HF part is then decomposed into a "rain component" and a "non-rain component" using sparse coding. The model experiments were conducted on synthetic rain images built using an image software, with the model achieving a VIF of 0.60. While the method has some degree of performance with common rain conditions, it has difficulties to handling more complex rain dynamic scenarios.

In <sup>[31]</sup>, an effective method based on simple patch-based priors for both the background and rain layers is proposed, which is based on the Gaussian mixture model (GMM) to accommodate multiple orientations and scales of the rain streaks. The two GMMs for the background and rain layers, defined as GB and GR, are based on a pre-trained GMM model with 200 mixture components. The method was evaluated using synthetic and real images, and the results compared to SOTA methods, with the proposed method achieving an SSIM of 0.88.

In <sup>[32]</sup>, it is proposed a DNN architecture called DerainNet for removing rain streaks from an image. The architecture is based on a CNN, enabling the direct map of the relationship between rainy and clean image detail layers from the data. For effective enhancement, each image is decomposed into a low-frequency base layer and a high-frequency detail layer. The detail layer corresponds to the input to the CNN for rain removal to be combined at a final stage with the low-frequency component. The CNN model was trained using synthesized images with rain, with the model achieving an SSIM of 0.900, increasing by 2% the performance in comparison to <sup>[31]</sup> using GMM.

A performance evaluation of the reviewed image de-raining methods is given in Table 3.

Deference	Main Tachniquaa	# Data	Drag/Cono
Relefence	Man rechniques	Success	Pros/Cons
[23]	GANS,	UCID	Robust,
	conditional GAN	PSNR 24.34	SOTA
[ <u>30]</u>	Bilateral filter,	Systeticg	Simple,

**Table 3.** Performance evaluation of the reviewed image de-raining methods.

		# Data	
Reference	Main Techniques		Pros/Cons
		Success	
	image decomposition	VIF 0.60	Parameter dependent
[31]	GMM,	Systetic	Simple,
	image decomposition	SSIM 0.880	Pre-trained dependent
[ <u>32</u> ]	CNN,	Systetic	Simple,
	HF component layer	SSIM 0.900	Robust

#### References

- 1. Wu, L.; Wang, Y.; Gao, J.; Li, X. Deep adaptive feature embedding with local sample distributions PSNR—peak signal to noise ratio (higher, the better; range: [0.0,—]), VIF—Information Fidelity (higher the better, for person re-identification. Pattern Recognit. 2018, 73, 275–288. range: [0.0, 1.0]), SSIM—structural similarity index (higher, the better; range: [0.0, 1.0]).
- 2. Zhang, W.; Ma, B.; Liu, K.; Huang, R. Video-based pedestrian re-identification by adaptive spatiotemporal appearance model. IEEE Trans. Image Process. 2017, 26, 2042–2054.
- Varior, R.R.; Haloi, M.; Wang, G. Gated Siamese Convolutional Neural Network Architecture for Human Re-Identification. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 791–808.
- Xiao, T.; Li, H.; Ouyang, W.; Wang, X. Learning deep feature representations with domain guided dropout for person re-identification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vegas, NV, USA, 27–30 June 2016; pp. 1249–1258.
- McLaughlin, N.; Martinez del Rincon, J.; Miller, P. Recurrent convolutional network for videobased person re-identification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vegas, NV, USA, 27–30 June 2016; pp. 1325–1334.
- Yan, Y.; Ni, B.; Song, Z.; Ma, C.; Yan, Y.; Yang, X. Person re-identification via recurrent feature aggregation. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 701–716.
- Yi, D.; Lei, Z.; Liao, S.; Li, S.Z. Deep metric learning for person re-identification. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Washington, DC, USA, 24–28 August 2014; pp. 34–39.

- 8. Zheng, Z.; Zheng, L.; Yang, Y. A discriminatively learned cnn embedding for person reidentification. ACM Trans. Multimed. Comput. Commun. Appl. (TOMM) 2017, 14, 1–20.
- Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
- 10. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 2012, 25, 1097–1105.
- LeCun, Y.; Boser, B.; Denker, J.; Henderson, D.; Howard, R.; Hubbard, W.; Jackel, L. Handwritten digit recognition with a back-propagation network. Adv. Neural Inf. Process. Syst. 1989, 2, 396– 404.
- 12. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
- 13. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv 2016, arXiv:1602.07261.
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- 15. Canziani, A.; Paszke, A.; Culurciello, E. An analysis of deep neural network models for practical applications. arXiv 2016, arXiv:1605.07678.
- 16. Gong, S.; Cristani, M.; Yan, S.; Loy, C.C.; Re-Identification, P. Springer Publishing Company. Incorporated 2014, 1447162951, 9781447162957.
- 17. Li, D.; Zhang, Z.; Chen, X.; Ling, H.; Huang, K. A richly annotated dataset for pedestrian attribute recognition. arXiv 2016, arXiv:1603.07054.
- Gray, D.; Tao, H. Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features. In Proceedings of the European Conference on Computer Vision, Marseille, France, 12–18 October 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 262–275.
- Lu, X.Y.; Skabardonis, A. Freeway traffic shockwave analysis: Exploring the NGSIM trajectory data. In Proceedings of the 86th Annual Meeting of the Transportation Research Board, Washington, DC, USA, 21–25 January 2007.
- 20. Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets robotics: The kitti dataset. Int. J. Robot. Res. 2013, 32, 1231–1237.
- 21. Everingham, M.; Van Gool, L.; Williams, C.K.; Winn, J.; Zisserman, A. The pascal visual object classes (voc) challenge. Int. J. Comput. Vis. 2010, 88, 303–338.

- 22. Lowe, D.G. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 2004, 60, 91–110.
- 23. Zhang, H.; Sindagi, V.; Patel, V.M. Image de-raining using a conditional generative adversarial network. IEEE Trans. Circ. Syst. Video Technol. 2019, 30, 3943–3956.
- Schaefer, G.; Stich, M. UCID: An uncompressed color image database. In Storage and Retrieval Methods and Applications for Multimedia 2004; International Society for Optics and Photonics: Washington, DC, USA, 2003; Volume 5307, pp. 472–480.
- 25. Arbelaez, P.; Maire, M.; Fowlkes, C.; Malik, J. Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2010, 33, 898–916.
- 26. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 2004, 13, 600–612.
- 27. Wang, Z.; Bovik, A.C. A universal image quality index. IEEE Signal Process. Lett. 2002, 9, 81-84.
- 28. Sheikh, H.R.; Bovik, A.C. Image information and visual quality. IEEE Trans. Image Process. 2006, 15, 430–444.
- Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the 32nd Conference on Neural Information Processing Systems, Montreal, QC, Canada, 7–10 December 2015; pp. 91–99.
- 30. Kang, L.W.; Lin, C.W.; Fu, Y.H. Automatic single-image-based rain streaks removal via image decomposition. IEEE Trans. Image Process. 2011, 21, 1742–1755.
- Li, Y.; Tan, R.T.; Guo, X.; Lu, J.; Brown, M.S. Rain streak removal using layer priors. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2736–2744.
- 32. Fu, X.; Huang, J.; Ding, X.; Liao, Y.; Paisley, J. Clearing the skies: A deep network architecture for single-image rain removal. IEEE Trans. Image Process. 2017, 26, 2944–2956.

Retrieved from https://encyclopedia.pub/entry/history/show/46376