OSJ | OPEN SCIENCE JOURNAL

REVIEW

Review of Current Methods for Re-Identification in Computer Vision

Matthew Millar[1]*

[1]Northcentral University, Japan

*Corresponding author: Matthew Millar: matthew.c.r.millar@gmail.com

Citation: Millar M. (2019) Review of Current Methods for Re-Identification in Computer Vision. Open Science Journal 4(1)

Received: 11th June 2019

Accepted: 5th August 2019

Published: 11th September 2019

Funding: The author(s) received no specific funding for this work

Competing Interests: The author have declared that no competing interests exists.

Abstract:

The problem of reidentification of a person in multiple cameras is a hot topic in computer vision research. The issue is with the consistent identification of a person in multiple cameras from different viewpoints and environmental conditions. Many computer vision researchers have been looking into methods that can improve the reidentification of people for many real-world purposes. There are new methods each year that expand and explore new concepts and improve the accuracy of reidentification. This paper will look at current developments, and past tends to find what has been done and what is being done to solve this problem. This paper will start by introducing the topic as well as covering the basic concepts of the reidentification problem. Next, it will cover standard datasets for today's research. Then it will look at evaluation techniques. Then this paper will start to describe simple techniques followed by the current deep learning techniques. This paper will cover the use of these techniques, what are some of their weaknesses and their strengths. It will conclude with an overview of some of the best models and show which models have the most promise and which models are not as useful.

Keywords: Computer Vision; Re-identification; ReID; Market1501; Tracking

Introduction

The problem of accurately locating a target in either multiple photographs, a video stream, or from a collection of cameras at different angles has been a hot topic in computer vision research. Re-identification (ReID) has many useful real-world applications. These applications can range from security systems and surveillance to automatically tagging people in photographs on social media. ReID can monitor a location and track a suspect as they traverse different areas. ReID can significantly reduce the time and cost in apprehending and finding a potential target. ReID can also be used to monitor customer's activity in a store to aid in streamlining the geospatial layout of a store's products. Combined with data analytics, ReID can be a potent tool for decision-makers and planners.

ReID is known as the process of locating an individual target in either an image or in video frames [1, 2, 3] over some time. The ReID model should accurately track and identify a target in real-time if possible. The process of ReID focuses on correctly matching a target from either a sample or a dataset and find the target in another image regardless of any change in position or location. These changes can range from altered lighting, the target's position or pose, the scale of the target, background changes, and camera rotations or orientation [4].

The difficulty of ReID

ReID faces many challenges to find a target in multiple images accurately. Camera calibration [5] or different types of cameras can increase the difficulty in ReID a person. A change in a camera by either; height, angle, brand, lens, and even placement can have adverse effects on subject identification [6, 4]. This issue is known as an intra-class variation. The change in lighting conditions can also affect some models. Low resolution is a pervasive issue in ReID as some sources of input come from older or low-quality cameras. Local feature extraction can be more difficult as there are little or no small details in an image from low-quality cameras. Occlusion can also increase the difficulty in ReID of targets, especially in video files or when there are many people in the scene. Uniforms or similar clothing can also increase the difficulty when not using facial recognition or when there is no unique feature per person to use. Scalability is another issue with ReID. The number of targets can be huge, the number of cameras, the amount of streaming footage, and the total number of people can add to the complexity.

Datasets

There are a few standard datasets that are used for training and testing ReID models. The most popular ones are Market-1501, CUHK01-03, and VIPeR [7, 8, 9, 10, 11]. Other datasets are PRW (Person Re-identification in the Wild), CAMPUS, RAiD, EPFL, and i-LIDS [12, 13, 14, 15, 16]. These datasets can be either image data or video files. The datasets typically consist of a training and testing set, but sometimes they have more testing and even partial or occluded image sets for evaluation purposes. These datasets have bounding boxes defined and, in most cases, pre-cropped images to aid in training.

Evaluation

The most common type of comparison found in the literature between each model is mean average precision (mAP), Rank 1, Rank 5, and Rank 10. The mAP is a standard evaluation tool for calculating the accuracy for object detection models. Rank X is the accuracy of finding the same person in the xth photo or frame [12, 17, 18]. These two metrics have been used in multiple articles for model comparison and are acting as the industry-standard metrics.

Simple Techniques Used

There have been many techniques used to perform ReID in images. These techniques can be simple as color feature extraction, shapes, gait analysis, or complex as multi-layer deep learning models that use local and global features for the analysis. The gait analysis is one of the least favorite methods for ReID because of the high possibility of the change in rotation, scale, and angle of view of the target. These alterations can make matching the gate of a person nearly impossible [4, 19]. Feature extraction based on the target's physical appearance can be some of the more straightforward methods to implement and give decent results depending on the environmental circumstances [20]. Color histograms can be used for looking at targets precisely and not that of a whole scene as this would lead to inaccurate results due to the changing of people, background, lighting, and position that would alter the color between each image [21, 22]. The use of color histograms is a simple method for photo similarity tests by looking at the distance between each histogram. However, color histograms are very limited and will not capture the features in great detail, which will result in several false positives and negatives.

Feature extraction and the comparison between two different images via distance metrics have started becoming the primary method for ReID. Euclidean distance equation, triplet loss, and other distance measurements [23, 24] are an excellent way to tell similarities between two images. The lower (or smaller) the distance between the two features, the more likely these features are of the same target. Bhattacharyya coefficients are also useful for distance measurements [25]. The Bhattacharyya coefficient can be useful when different; extracted features fall outside the distance measurement space [4]. However, simple feature extraction cannot account for a total complexity of ReID. Feature extraction is a method for downsampling image data into either features or vectors for use in more complex neural networks.

Deep Learning for ReID

Color histograms, gait analysis, and simple distance measurement techniques are useful but can lack accuracy and dependability in real-world applications. The use of deep learning (DL) can overcome some of the shortcomings of the previously mentioned methods. DL has proven to be useful in many filed in computer vision and has shown in research to be very useful for ReID projects. Some of these DL models have to potential to exceed human-level accuracy [26]. These models can take many different forms, from a Siamese neural network [27] to more complex multilevel triplet DL models [28].

The first model approach is the simple Convolutional Neural Network (CNN), which is a common approach for all image problems. CNN for ReID problems should have two basic needs. One is a feature exaction method for getting a useful feature from an image, and the second one is a distance metric for the comparison between the similarities of two images [29, 30]. Some of the more accurate CNN approaches look at both global and local features when they are performing feature extraction [28, 10]. By adding both the global and local features of an image, the accuracy and reliability of the model seem to increase significantly when the combination of these features [31, 32]. The use of CNN alone does not yield state-of-the-art results, so the use of either global, local, or a combination of the two tend to yields results that are far more accurate than a CNN alone. Some research shows that the global features are more critical than local features, but the use of local features as a supplement can increase accuracy [10]. The local features should not be in the inference section of the model. However, the local features should aid in the global feature learning of a model. This method may reduce the added complexity of combining local feature, a global feature, and even other models feature extraction together [10]. Thus, reducing the dimension of the image significantly without losing valuable information about the target.

Siamese Convolutional Neural Networks (S-CNN) have given good results for ReID [27]. S-CNN can learn embedded features of both similar and dissimilar pairs that belong to a group of images. The 'margin' is the distance that separates the two features [27]. S-CNN does not capture the link between a group of images as it only extracts a fixed feature for a single image and does not update the extracted features for correctly paired images, which does not allow for the model to update its patterns, and can hinder future matches [27, 33, 34]. However, S-CNN can produce excellent results reliably with lower complexity than other models that rely on multiple layer feature extraction.

Deep metric learning (DML) has also seen use in ReID. DML works by creating an embedding feature from an image and checking the distance between each feature. There are two pairs of positive and negative images. Triplet loss is a typical distance measurement between the two sets of pairs [35]. The proper selection of samples for the triplet loss function is vital for the success of this method. There should be both an easy pair and a difficult pair when training the model. The easy pair should have a small distance or a slight change between the two images. Changes can be in the rotation of an image or other small changes. The hard pair would be a more significant change in either clothing, surroundings, lighting, or other drastic changes. Doing this can significantly improve the accuracy of the triplet loss function and at the same time, the accuracy of the DML system [36, 37].

Re-Ranking is a process for using the L2 Euclidean distance combined with Jaccard distance metrics to find similar images [38]. Re-Ranking has shown to be successful for object retrieval, especially when with the k-nearest neighbor distance algorithm [39]. Re-Ranking uses the distance between the query image, positive image, and negative image to decide the similarity between them. Feature representation or metric learning are two of the primary methods used for ReID in Re-Ranking [40, 41]. Re-Ranking has proven to be a simple yet effective means for ReID.

Human joints and human key points can be used instead of features extracted from an image as in the more common CNN approaches. An image pool where each target has multiple pictures stored of them is a common approach for key point extraction and comparison. The human joint comparison will need more than three images for comparison as it does not use a negative image. This approach will need multiple images on one target to learn how to map joint data to the image. The rules of the image pool can be adjusted to increase the representation of each target. The ranking of each sample in the image pool can be calculated using a time series model by finding the distance between each joint. Then a re-ranking algorithm can be used to adjust the initial ranking list, which is in ascending order [42]. Re-Ranking has shown to provide results that are comparative to that of other state-of-the-art CNN models. Human pose estimation can also be a solution for ReID, but this approach will suffer as humans are flexible and dynamic and will not maintain the same position for very long. So, the pose models results will suffer from either need a considerable amount of data or an overly complicated model that underperforms.

Features extraction from a target to increase the accuracy of ReID has been shown to significantly improve the accuracy of ReID [4, 10, 26]. A unique method for using said features is Boosting Ranking Ensemble [43], which uses the ranking method to create the features from the images. This method combines multiple features and metrics which can then be used to give even better results than current feature extraction methods. The use of sectioning attributes of a person's appearance is useful as a means to extract a feature or multiple features. By defining a person by how they look or what they are wearing, allows for different approaches to be used for ReID. A CNN model to classify the items that a person is wearing [47] is one method. The use of SVM to create the lower-level descriptors, which is useful in a metric learning model for ReID [48] is another method for attribute defining. Attribute defining has unique ReID methods as it can be used to identify a person if their identity is known previously [47]. This method will allow for accurate tracking of an individual over several cameras.

Unsupervised learning techniques have also proven to be very useful in ReID and sometimes have proven to be very accurate, but not as accurate as supervised learning. One of the more unique approaches which did quite well as a method that created a person tracklet automatically and did not need a pairwise ID label for the data. This method used a local feature as the label discrimination, which would allow for ReID between different cameras. This local feature would be created for each camera separately. Then a global cross camera tracklet matching would be made by grouping images that are similar based on the local features that for each camera [44]. The use of small subsets of a person that is in a dictionary of sparse atoms has also been used to get a representation of a person for ReID [45]. The use of a dynamic graph matching that can drastically improve label generation by dynamically adjusting the structure of the graph. These adjustments are made by a positive re-weighting method to help manage noisy data issues found in real-world data [46]. A method for multi-labeling using mainly hard negatives in a triplet loss method can help improve results for unsupervised learning approaches. A similar pair is two images that have similar characteristics, and if they do not, then it can be classified as a hard-negative pair [49]. A tight threshold would control how similar the characteristics are.

Methodology Comparison

The result from the current research still shows sign of improvement in ReID. By comparing mAP results from different studies, it is clear to see that the best model is challenging to determine.

Table 1 The outcome of different studies

Methodology/Model	mAP
AlignedReID[10] DL+ Global Feature + Local Feature	90.7
GLAD[50] DL	73.9
In(RK)[37] DL + Triple Loss	81.1
Deep[51] DL	68.8
APR[17] DL + Attribute Data	66.89
S-CNN+Gate[27] DL + S-CNN Framework + Gate	48.45
TAUDL[32] DL (Unsupervised)	31.2
St-ReID[52] DL + Spatial + Temporal	**95.5**

Table 1 shows that ReID has more room for improvement. Even models that use DL rarely get over 90 in the mAP scores. The use of DL does increase accuracy over other simpler models; however, it is not enough to capture the complicated relationship between two similar images fully. Using other complex algorithms, more complex feature extraction techniques, multiple level features, and a combination of other metrics can significantly improve ReID.

Conclusion

ReID has many real-world applications but is a very complicated task, especially in the wild. To successfully find a person from one camera to another camera is only made more difficult as the locations and environment change. Simple techniques can be used to a degree of success but gives poor results for more complex changes of environment and other factors that affect computer vision. The use of DL models can significantly improve the quality and accuracy of attempts as solving ReID. DL models can overcome the majority of the issues that face the simpler models. By combining both a local and global feature extraction, these DL models can become even more accurate and can then be used to identify people from different cameras successfully. Unsupervised methods are an excellent way to deal with ReID when looking at a static environment, but supervised learning has shown to be a better choice when it comes to the overall accuracy and when the cameras positions and environment changes.

Future Works

Further research is needed when looking at improving the accuracy of ReID task. The improvement should come in the form of deeper networks of combined multi-level feature extraction and the use of triple loss function. The triple loss function has proven very beneficial for similarity comparisons of different models in the past [28, 36, 37]. By using triple loss, the more similar images will be closer together, while different images will be further apart. Triple loss is far more efficient at defining the similarity between images over the cross-categorical loss. With the triple loss, the neural network does not have to be built for classification, but instead feature extraction, which is the primary output for ReID projects.

The use of multi-level features can also give better results over the more typical global features only. Global features can give accuracy up to 70-80% [26, 4], but the combination of local level and global level features have shown to improve performance of ReID algorithms significantly to much higher than 90% [10]. Few studies have looked at combining the different level of feature extraction or using multi-level feature extraction to augment other levels. Thus, multi-level features are an excellent spot to improve the accuracy and increase the body of knowledge in the ReID task. The combination of spatial and temporal features [52] have also shown great promises. A possible next step would be to combine a mixture of local and global features on a temporal Long Term Short Memory (LSTM) CNN, which may yield the best possible results. However, this would have to be done on a particular dataset, as the current dataset lacks temporal labels so it would be very difficult to determine which image came first and which was last.

References:

[1] Allouch, A. 2010. People re-identification across a camera network. Master's thesis in computer science at the school of computer science and engineering royal institute of technology, Stockholm, Sweden,

[2] Cong, T., Achard, C., Khoudour, L. and Douadi, L. 2009. Video sequences association for people re-identification across multiple non-overlapping cameras. Proc. Int. Conf. Image Analysis and Processing (ICIAP), Vietri Sul Mare, Italy, 179–189. DOI: https://doi.org/10.1007/978-3-642-04146-4_21

[3] Guermazi, R., Hammami, M. and Hama, A.B. 2009. Violent web images classification based on MPEG7 color descriptors. Proc. IEEE Int. Conf. Systems Man and Cybernetics, San Antonio, USA, October 2009, 3106–3111. DOI: 10.1109/ICSMC.2009.5346149

[4] Saghafi, M. A., Zaman, H. B., Saad, M. H., and Hussain, A. 2014. Review of person re-identification techniques. IET Computer Vision,8(6), 455-474. DOI:10.1049/iet-cvi.2013.0180

[5] M. J. Lyons, J. Budynek and S. Akamatsu, 1999. "Automatic classification of single facial images," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 12, 1357-1362. DOI: 10.1109/34.817413

[6] Zhao, C., Chen, K., Wei, Z., Chen, Y., Miao, D., and Wang, W. 2019. Multilevel triplet deep learning model for person re-identification. Pattern Recognition Letters. 117, 161-168, DOI.org/10.1016/j.patrec.2018.04.029.

[7] Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., and Tian, Q. 2015. Scalable Person Re-identification: A Benchmark. 2015 IEEE International Conference on Computer Vision (ICCV). DOI:10.1109/iccv.2015.133

[8] Li, W., Zhao, R., Xiao, T., and Wang, X. 2014. DeepReID: Deep Filter Pairing Neural Network for Person Re-identification. 2014 IEEE Conference on Computer Vision and Pattern Recognition. DOI:10.1109/cvpr.2014.27

[9] Li, W., Zhao, R., and Wang, X. 2013. Human Reidentification with Transferred Metric Learning. Computer Vision – ACCV 2012 Lecture Notes in Computer Science, 31-44. DOI:10.1007/978-3-642-37331-2_3

[10] Li, W., and Wang, X. 2013. Locally Aligned Feature Transforms across Views. 2013 IEEE Conference on Computer Vision and Pattern Recognition. DOI:10.1109/cvpr.2013.461

[11] Van Beeck K., Van Engeland K., Vennekens J., and Goedemé T. 2017. Abnormal Behavior Detection in LWIR Surveillance of Railway Platforms. Proceedings of the 14th IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS). DOI: 10.1109/AVSS.2017.8078540

[12] Zheng, L., Zhang, H., Sun, S., Chandraker, M., Yang, Y., and Tian, Q. 2017. Person Re-identification in the Wild. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). DOI:10.1109/cvpr.2017.357

[13] Geiger, A., Lenz, P., Stiller, C., and Urtasun, R. 2013. Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research, 32(11), 1231-1237. DOI:10.1177/0278364913491297

[14] Das, Abir & Chakraborty, Anirban and K. Roy-Chowdhury, Amit. 2014. Consistent Re-identification in a Camera Network. DOI: 10.1007/978-3-319-10605-2_22.

[15] Chavdarova, T., and Fleuret, F. (2017) Deep Multi-Camera People Detection. Proceedings of the IEEE International Conference on Machine Learning and Applications. arXiv:1702.04593v3 [cs.CV] Retrieved from https://arxiv.org/abs/1702.04593

[16] Li, Y., Wu, Z., Karanam, S., and Radke, R. 2015. Multi-Shot Human Re-Identification Using Adaptive Fisher Discriminant Analysis. 73.1-73.12. DOI:10.5244/C.29.73.

[17] Lin, Y., Zheng, L., Zheng, Z., Wu, Y., and Yang, Yi. (2017). Improving Person Re-identification by Attribute and Identity Learning. arXiv:1703.07220v2 [cs.CV] Retrieved from https://arxiv.org/abs/1703.07220

[18] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian. 2015. Scalable person re-identification: A benchmark. ICCV.

[19] Hamdoun, O., Moutarde, F., Stanciulescu, B., and Steux, B. 2008. Interest points harvesting in video sequences for efficient person identification. Proc. Eighth Int. Workshop on Visual Surveillance, Marseille, France.

[20] Zeng, G., Hu, H., Geng, Y., and Zhang, C. 2014. A person re-identification algorithm based on color topology, 2014 IEEE International Conference on Image Processing (ICIP), Paris, 2447-2451. DOI: 10.1109/ICIP.2014.7025495

[21] Farenzena, M., Bazzani, L., Perina, A., Murino, V., and Cristani, M (2010) Person re-identification by symmetry-driven accumulation of local features. Proc. Int. Conf. Computer Vision and Pattern Recognition (CVPR) 2360–2367

[22] Gheissari, N., Sebastian, T.B., Tu, P.H., Rittscher, J., and Hartley, R. 2006. Person re-identification using spatiotemporal appearance. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). 1528–1535. DOI: 10.1109/CVPR.2006.223

[23] Schroff, Florian, Dmitry Kalenichenko, and James Philbin. 2105. Facenet: A unified embedding for face recognition and clustering. Proceedings of the IEEE conference on computer vision and pattern recognition. DOI: 10.1109/CVPR.2015.7298682

[24] Luiten, J., Voigtlaender, P., and Leibe, B. 2019. PReMVOS: Proposal-Generation, Refinement and Merging for Video Object Segmentation. Computer Vision – ACCV 2018 Lecture Notes in Computer Science, 565-580. DOI:10.1007/978-3-030-20870-7_35

[25] D'angelo, A., and Dugelay, J.L. 2011. People re-identification in camera networks based on probabilistic color histograms. Proc. Electronic Imaging Conf. 3D Image Processing (3DIP) and Applications, 23–27. DOI: 10.1117/12.876453

[26] Xuan Zhang et al. AlignedReID: Surpassing Human-Level Performance in Person Re-Identification (2018). https://arxiv.org/abs/1711.08184 Retrieved from https://arxiv.org/abs/1711.08184

[27] Rama Varior, Rahul & Haloi, Mrinal and Wang, Gang. 2016. Gated Siamese Convolutional Neural Network Architecture for Human Re-identification. 9912. 791-808. arXiv:1607.08378 [cs.CV] Retrieved from https://arxiv.org/abs/1607.08378

[28] Cheng, D., Gong, Y., Zhou, S., Wang, J., and Zheng, N. 2016. Person Re-identification by Multi-Channel Parts-Based CNN with Improved Triplet Loss Function. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). DOI:10.1109/cvpr.2016.149

[29] Krizhevsky, A., Sutskever, I., and Geoffrey, E. H. 2012. ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems. 25. 10.1145/3065386.

[30] Girshick, R.B., Donahue, J., Darrell, T., and Malik, J. 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 580-587.

[31] Marchwica, P., Jamieson, M., and Siva, P. 2018. An Evaluation of Deep CNN Baselines for Scene-Independent Person Re-identification. 2018 15th Conference on Computer and Robot Vision (CRV). DOI:10.1109/crv.2018.00049

[32] Fan, H., Zheng, L., Yan, C., and Yang, Y. 2018. Unsupervised Person Re-identification. ACM Transactions on Multimedia Computing, Communications, and Applications,14(4), 1-18. DOI:10.1145/3243316

[33] Droghini, D., Vesperini, F., Principi, E., Squartini, S., and Piazza, F. 2018. Few-Shot Siamese Neural Networks Employing Audio Features for Human-Fall Detection. Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence - PRAI 2018. DOI:10.1145/3243250.3243268

[34] Chung, D., Tahboub, K., and Delp, E. J. 2017. A Two Stream Siamese Convolutional Neural Network for Person Re-identification. 2017 IEEE International Conference on Computer Vision (ICCV). DOI:10.1109/iccv.2017.218

[35] Liu, H., Feng, J., Qi, M., Jiang, J., and Yan, S.. 2017. End-to-end comparative attention networks for person re-identification. IEEE Transactions on Image Processing.

[36] Chen, W., Chen, X., Zhang, J., and Huang, K..2017. Beyond triplet loss: a deep quadruplet network for person re-identification. arXiv:1704.01719 [cs.CV] Retrived from https://arxiv.org/abs/1704.01719

[37] Hermans, A., Beyer, L., and Leibe, B.. 2017. In defense of the triplet loss for person re-identification. arXiv:1703.07737 [cs.CV] Retrieved from https://arxiv.org/abs/1703.07737

[38] Z. Zhong, L. Zheng, D. Cao, and S. Li. Re-ranking person re-identification with k-reciprocal encoding. arXiv:1701.08398 [cs.CV] Retrieved from https://arxiv.org/abs/1701.08398

[39] Chum, O., Philbin, J., Sivic , J., Isard, M., and Zisserman, A. 2007. Total recall: Automatic query expansion with a generative feature model for object retrieval. ICCV. DOI: 10.1109/ICCV.2007.4408891

[40] Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., and Tian, Q. 2015. Scalable person re-identification: A benchmark. ICCV. DOI: 10.1109/ICCV.2015.133

[41] Gray, D. and Tao, H. 2008. Viewpoint invariant pedestrian recognition with an ensemble of localized features. ECCV. DOI https://doi.org/10.1007/978-3-540-88682-2_21

[42] Yuan, M., Yin, D., Ding, J., Zhou, Z., Zhu, C., Zhang, R. and Wang, A. 2019. A multi-image Joint Re-ranking framework with updateable Image Pool for person re-identification. Journal of Visual Communication and Image Representation,59, 527-536, DOI: https://doi.org/10.1016/j.jvcir.2019.01.041.

[43] Li, Z., Han, Z., Xing, J., Ye, Q., Yu, X., and Jiao, J. 2019. High performance person re-identification via a boosting ranking ensemble. Pattern Recognition, 94, 187-195, DOI:https://doi.org/10.1016/j.patcog.2019.05.022.

[44] Li, M., Zhu, X., and Gong, S. 2019. Unsupervised Tracklet Person Re-Identification. IEEE Transactions on Pattern Analysis and Machine Intelligence. doi: 10.1109/TPAMI.2019.2903058

[45] Lisanti, G., Martinel, N., Micheloni, C., Bimbo, A. D., and Foresti, G. L. 2019. From person to group re-identification via unsupervised transfer of sparse features. Image and Vision Computing,83-84, 29-38. DOI:10.1016/j.imavis.2019.02.009

[46] Ye, M., Li, J., Ma, A. J., Zheng, L., and Yuen, P. C. 2019. Dynamic Graph Co-Matching for Unsupervised Video-Based Person Re-Identification. IEEE Transactions on Image Processing,28(6), 2976-2990. DOI:10.1109/tip.2019.2893066

[47] Lin, Y., Zheng, L., Zheng, Z., Wu, Y., Hu, Z., Yan, C., & Yang, Y. 2019. Improving Person Re-identification by Attribute and Identity Learning. Pattern Recognition. DOI:10.1016/j.patcog.2019.06.006

[48] Layne, R., Hospedales, T., & Gong, S. (2014). Re-id: Hunting Attributes in the Wild. Proceedings of the British Machine Vision Conference 2014. doi:10.5244/c.28.1

[49] Yu, H., Zheng, W., Wu, A., Guo, X., Gong, S., Lai, J. 2019. Unsupervised Person Re-identification by Soft Multilabel Learning.

arXiv:1903.06325 [cs.CV]. Retrieved From https://arxiv.org/abs/1903.06325