IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727 PP 18-23
www.iosrjournals.org

Person Re-Identification – A Mini Survey Report

Dr.S. Gomathi Meena[1], *Assistant Professor, Department of Computer Science, Vidhya Sagar Women's College*
P. Sathya Priya[2], *Student, B.Sc(Computer Science), Department of Computer Science, Vidhya Sagar Women'sCollege,*
S.Hemavathi Selvam[3], *Student, B.Sc(Computer Science), Department of Computer Science, Vidhya SagarWomen's College,*

Abstract: *Man or woman search has come to be a prime field due to its need in network and inside the area of research among researchers. These challenge objectives to find a probe character from whole scene which indicates super importance in video surveillance field to long lost humans, re-identity, and verification of person. In remaining few years, deep studying has played unremarkable function for the solution of re-identification trouble. Deep studying indicates high-quality performance in person (Re-ID) and search. Researchers experience more flexibility in presenting new methods and solve challenging issues such as low decision, pose variant, history clutter, occlusion, viewpoints, and low illumination. In particular, Convolutional Neural Network (CNN) achieves breakthrough performance and extracts useful patterns and features. Developing a new framework requires considerable effort. It takes a lot of effort and computational cost to get good results. This overview document contains a brief description of Feature Representation Learning and Deep Metric Learning and the new loss function. We are carefully examining the datasets along with performance analysis of the existing datasets. Finally, we review the current solution for further consideration.*
Keywords: *Person Re-Identification, Privacy, Person search, Metric learning*

I. Introduction

Person Re-ID has recently attracted the attention of scientists in the field of computer vision. Demand for reliable and intelligent video surveillance is increasing in today's society due to the importance of security purposes such as crime and terrorism prevention and forensic investigation. As an important and stressful hassle in pc vision, individual re-identity[1,2] and individual seek have emerged an impartial subject matter and rapid-developing topic in laptop imaginative and prescient that offers with individual retrieval in motion pictures and virtual images[3]. Deep learning has become a mainstream method for researchers. The triumph of deep learning methods has led to a new wave of people re-identification, pushing toward an exploratory reversal. Identification (re-identification) is widely used in academia and public safety, in large-scale industrial projects such as people tracking, behavioural analysis and surveillance in a wide range of public parks, universities and streets. From the standpoint of the video surveillance community, the first and foremost problem of human finding is the exact matching of human testers observed by cameras in different locations with intense changes in posture, viewpoint, and lighting. One way is to match the person under study to a gallery of hand-carved ones that differ slightly from the actual application is shown in ***Fig 1***.

Fig 1. Person Re-Identification – a simple view.

However, because human detection is close to the physical world and quite challenging, automated pedestrian detection gives false detections and disproportionate images. Here's why to solving the complex task of person Re-ID and person retrieval is very difficult due to the person's posture, lighting, video camera positioning, low resolution, occlusion, perspective, posture evaluation, and many image variations[5–10]. Low-resolution cameras cannot detect and track biometric features such as faces and gait patterns. The Re-ID and person search task only rely on visual appearance[13]. This document provides a broad and in-depth overview of people search. The purpose of the survey is to comprehensively develop the six parts of people search methods, including feature learning, architecture design, metric deep learning, and loss function design. In addition to the taxonomic discussion of previous people search methods, we take a closer look at the previous people search dataset. We comprehensively look at the effectiveness of people-finding methods and present promising future work areas today.

II. Related work

The history of finding people is relatively small. Recent work on people search is mainly related to people search using hand-crafted features and deep learning methods on images [4– 11]. Upon the rapid progression of CNN, various deep learning-based person search approaches have emerged. However, mostly surveys and researches emphasizing hardly on person (Re-ID). We have accurately and carefully selected remarkable and outstanding papers published in outstanding journals and major conferences. This review focuses on the rapid development of body detection over the past four to six years. Apart from this, other related works are also included in to make our work in this field clearer and more useful. We limit this review paper to a small discussion (re-identification) of the person retrieval method and recent individual work. Also, some work related to identity (re-identification) will be included in the discussion. We thoroughly reviewed many deep models, including identification, validation, deep metric learning, video-based deep models, and data augmentation. He also reviewed previous figures (re-identified), datasets in detail over the years, and discussed future ideas. From a people finding (re-identification) perspective, these surveys discuss and do not focus on people finding. We present a structured and comprehensive overview of deep learning algorithms that mutually process human detection and re-identification. This review is based on a critical and in-depth analysis of people search.

2.1 Summarizing points

A systematic review of how to find people of their decouple and with generalize existing people search methods in several aspects, including feature learning, multi-scale feature learning, architecture design, loss function, and deep metric learning. The proposed taxonomy is intended to help new researchers gain a comprehensive and deep understanding of people search.

A comprehensive review and analysis of the effectiveness of people search. We evaluate based classification of people search approaches and analyze the results of these surprising methods on people (re-identification) in existing datasets.

Here mentioned some important points for the development of the human search algorithm and offered future ideas and directions to the researchers.

III. Person Re-ID system approaches

Image-based Re-ID has received more attention in Human Re-ID research. It focuses on juxtaposing static images of people from non-overlapping camera views. Shape-based features such as clothing are commonly used for this task. Image-based Human Re-ID investigated different aspects of different scenarios to improve the results. One such aspect is the division of image-based human Re-ID models into three types based on image type: RGB, RGB-Depth and RGB-Infrared images. This approach required no knowledge of the camera's spatial distribution or assumptions about how people would move in the target environment. In [16] it implements a lightweight deep CNN architecture by generating new blocks of residuals combined with multiple convolutional function streams, each stream discovering a function in a specific way. A unified aggregation network was then developed to dynamically combine the multi-scale features with the weights associated with the input channels. Hong Xing et al [17] we built an unsupervised approach using an asymmetric human deep metric learning technique (re- identification) in an unsupervised manner. The specific projection view of each image is derivedfrom the asymmetric clustering of crosswalk images.

The model identifies shared regions where perspective bias is reduced and consequently improved matching performance can be achieved. The specific projection view of each image is derived from the asymmetric clustering of crosswalk images. The model defines shared regions where perspective bias can be reduced and consequently matching performance can be improved. New Deep Associative Learning [18] is a deep learning approach to unsupervised video (re-identification) extracted from untagged videos (Re-ID) in surveillance data. Long et al. [19] developed a framework for tracking multiple people and built a scoring function based on convolutional neural networks. Ihan et al. [20] developed a novel architecture for video-based human re-identification that relies on two components: a refinement module and a spatial cue integration module. Hongxing et al. [21] proposed a new DECAMEL framework that introduces an unsupervised loss function and includes an end-to-end learning function to solve the asymmetric metric learning problem. Another PS (Personal Search) approach, as opposed to re-identifying the person, detects all people in the image except for sentences and bounding boxes and recognizes the probe person is shown in ***Fig 2***.

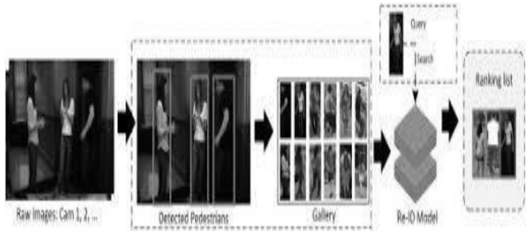

Fig 2. Overview of process for searching the person.

3.1 Learning of feature representation

We present several approaches to the study of symbols and several articles of people (re-identification) adopted in the human search structure to better understand the study of symbols. There are two main categories containing global and local features. The global features are extracted from the entire scene without isolating the additional signals. Local features divide the entire image into partial or local features to create a collective representation of each image.

3.1.1 Global feature of representation learning

Zhong et al. [22] used the VGG-16 network to derive 512-dimensional feature vectors from human images. A pair of features is assigned to each node, the edges are also in the same context, and the global part is connected to the graph to determine whether two input images belong to the same ID belong to the same ID. Expanded separate characteristics with top result matched context sets, and all these features are exhibited using contextual graph. Cunyuan et al. [23] introduced non-local components that allows model to learn more global features, focus on region in picture where scene is crowded.

3.2 An approach for deep metric learning on RGB

Deep metric learning is a method for calculating the distance between two data spaces and determining whether objects are similar. The deep metric learning approach is to increase the distance between heterogeneous data spaces and decrease the distance between similar objects [24,25].

The first layer is a convolutional layer, which was used to extract local features from the image. We then adopted a maximum pooling layer of to make features above robust against local bias. Subsequently, patches corresponding to layer were added to match the output of the local patch filter across the different representations. In order to improve the reliability of patch matching, a maxout-grouping layer is adopted to divide the patch bias matrices into several groups, and only select the maximum activation values of from each group to form a single output of to form the next layer.

Filter pairs and maxout-grouping layers were used to study light intensity conversion. Geometric transformations were studied using patch matching layers, convolutional, maximum union layers, and fully connected layers. Finally, we took a fully connected layer and a softmax loss function with no measure whether the images of the two people are similar.

3.2.1 Video-based person Re-ID

Footage is more realistic for human re-identification. Video-based person re- identification matches a person video sequence from one camera to a gallery of video sequences recorded from other, non-overlapping cameras. Re-identification of people by video has not received as much attention as re-identification based on image. Re-identifying people using video has many advantages over static images. It provides successive images that contain information about the appearance of the same person, which is effective in reducing the impact of some ambiguous situations.

The first human Re-ID video model using deep learning was proposed by the authors in [26]. They presented a circular convolutional network that uses both the colour and optical features of the flow, where colour describes a person's appearance and optical flow describes short-term motion, including human gait and other motion signatures. Using a combination of colour and optical flow, the model should be able to better exploit the short-term properties to improve re-identification accuracy. One reliable function in the field of personal identification is biometrics. They fall into two main categories: basic biometrics such as fingerprints and faces, and soft biometrics such as gait. While the first category is incompatible with solving the human re-identification problem, “soft” biometrics, particularly ambulatory biometrics, are receiving considerable interest in video surveillance systems because of their multiple advantages. Soft biometrics can be extracted from low- resolution video and soft biometrics, such as gait, can be recorded with a camera at long distances [27]. They proposed handcrafted method for enhancing person Re-ID by integrating gait features with appearance features, also authors in [28] proposed handcrafted method for long-term person Re-ID using true motion features. The motion patterns were extracted by encoding trajectory-aligned descriptors with Fisher vectors in a spatial-aligned pyramid. This work supports long-term person Re-ID since it didn’t depend on appearance features. None of the existing long-term person Re-ID works was built based on the deep learning methods and this is one of the future directions in person Re-ID task.

IV. Benchmark dataset for Person Re-ID

Because it requires individual authority in an uncontrolled environment, generating datasets with repeated individual identities is a significant task and a security risk. Despite its importance, the search has garnered significant market demand. This is primarily because the process of re-identifying the person (Re-ID) is easier. First, training deep learning models requires large amounts of real-time data. To evaluate the robustness of human re- identification systems, it is important to train deep learning models and have data-rich human re-identification datasets with the characteristics of inter- and intra-class variation. These datasets differ in number of images, IDs, cameras, and image types. This dataset contains about 9,000 IDs and nearly 100,000 bounding boxes, providing a significant amount of data. The scientific community needs larger data sets to achieve good accuracy. Second, detecting human candidates in an image is always challenging, such as low-resolution issues, occlusion, offset, etc.

Cumulative Matching Characteristics (CMC) are used to evaluate people search and re-identification methods. CMC represents the probability of the requested image appearing in the candidate list due to the size difference. Given a valid image, only the first image is considered in the CMC calculation. So CMC is the correct method when actual request data exists. This criterion is used when researchers are more interested in bringing the Fundamental Truths (GTs) back to the top of the rankings.

4.1 Metrics of Re-ID evaluation

There are a few facts in the gallery, but Average Accuracy (mAP) is the best estimate for this scenario. mAP is used when two cameras capture the same GT but have unusual extraction capabilities. At this point, mAP works well instead of CMC because CMC has weak discriminatory power. In addition to the task of finding people, a candidate field is considered positive if its GT is greater than 0.5. A number of later research papers 62, 60, 3, 56, 57, 58, 63, 64] reported multi-significance mAP results. CUHK-SYSU contains 18,184 images, 96,143 pedestrians, ignoring background pedestrians with a height of N50 pixels.

V. Issues and Challenges

Most existing human re-identification models support short-term re-identification, where a person always moves in a small space for a short period of time, and the probe image of a person and the corresponding image in the gallery have a similar shape, such as clothes. However, the most realistic scenario would typically require long-term human re- identification systems, in which people are likely to emerge after a few days [29]. Soft biometrics are used with deep learning in the yield of human identification such as in [30,31]. Future person Re-ID datasets should consider the long-term datasets that recorded over several days [31,33]. Long-term deep person Re-ID becomes one of the main research directions.

Human attributes are high-level semantic representations, such as hair, gender, age, etc., and are resistant to many environmental changes. Several studies have adapted these properties to deep learning-based re-identification systems, such as the one in [34,35], to bridge the gap between images and high-level semantic features. Investigation of the characteristics of yielded the promised results, making it one of future destinations.

Search for people in large image datasets using natural language descriptions, not based on images, videos, or attributes. Re-identifying the person and extracting the most similar person image from the gallery is a useful task in many situations, such as when a probe image cannot be found. One of these works was done by [35]. The language-based Human Re-ID and Large Human Re-ID Description Dataset should be considered in a study of new Re-ID people.

VI. Conclusion

People search is a new paradigm that can be applied practically to large communities of people without spending a lot of time manually cropping images of peoples. This overview document describes recent work on locating person. We first discussed the history of people finding and the trend in recent years. Second, the process of searching and re-identifying individuals is discussed in detail with a literature review. Third, we reviewed the existing CNN architectures used for person retrieval to achieve this. The highly cited people search architecture has been described with a schematic explanation. Finally, we also provided a metric learning solution for people search. Human search, from the perspective of video surveillance systems, has increased interest in the computer vision industry. Therefore, we focus our development work on human search. A review of human search algorithms and datasets helps explore re-identification algorithms becomes collectively promote detection personnel and re-identification work. This should be done in the future to make it more reliable and accurate. However, although has improved using a deep learning method comparable to the manual method, there are still many issues and limitations that need to be considered in future research directions to improve the human re-identification system and achieve promising results.

References

[1]. Z. Liang, H. Zhang, S. Sun, M. Chandraker, Y. Yang, Q. Tian. Person re-identification in the wild, Proc. IEEE Conf. Comput. Vis. *Pattern Recognit.* (2017), pp. 1367–1376.

[2]. L. Zheng, Y. Yang, A. G. Hauptmann. Person re-identification: Past, present and future, arXiv Preprint, 2016 arXiv:1610.02984.

[3]. S. Zhai, S. Liu, X.Wang, J. Tang, FMT: fusingmulti-task convolutional neural network for person search, *Multimed. Tools Appl.* (2019), pp. 1–12.

[4]. T. Xiao, S. Li, B. Wang, L. Lin, Xi Wang. Joint detection and identification feature learning for person search, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017), pp. 3415–3424.

[5]. X. Chu, W. Ouyang, H. Li, X. Wang, Structured feature learning for pose estimation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016), pp. 4715–4723.

[6]. Wei Li, Rui Zhao, Tong Xiao, Xiaogang Wang. Deepreid: Deep filter pairing neural network for person re-identification, Proceedings of the IEEE conference on computer visionand pattern recognition (2014), pp. 152–159.

[7]. Shengcai Liao, Yang Hu, Xiangyu Zhu, Z.Li. Stan. Person re-identification by local maximal occurrence representation and metric learning, Proceedings of the IEEE Conferenceon Computer Vision and Pattern Recognition (2015), pp. 2197–2206.

[8]. Sakrapee Paisitkriangkrai, Chunhua Shen, Anton Van Den Hengel. Learning to rank in person re-identification with metric ensembles, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015), pp. 1846–1855.

[9]. Tong Xiao, Hongsheng Li, Wanli Ouyang, Xiaogang Wang. Learning deep feature representations with domain guided dropout for person re-identification, Proceedings of the IEEE conference on computer vision and pattern recognition (2016), pp. 1249–1258.

[10]. Liang Zheng, Liyue Shen, Tian Lu, Shengjin Wang, Jingdong Wang, Qi Tian. Scalable person re-identification: A benchmark, Proceedings of the IEEE international conference on computer vision (2015), pp. 1116–1124.

[11]. X. Yuanlu,M. Bingpeng, R. Huang, L. Lin. Person Search In A Scene By Jointly Modeling People Commonness And Person Uniqueness, Proceedings Of The 22nd ACM International Conference On Multimedia November 03–07 (2014), pp. 937–940.

[12]. S. Gong, M. Cristani, C.C. Loy, M.T. Hospedales, Person Re-Identification (Advances in Computer Vision and Pattern Recognition) (2014), pp. 301–313.

[13]. Qingming Leng, Mang Ye, Qi Tian. A survey of open-world person re-identification, IEEE Transactions on Circuits and Systems for Video Technology 30(4) (April 2020), pp. 1092–1108.

[14]. DiWu, Si-Jia Zheng, Xiao-Ping Zhang, Chang-An Yuan, Fei Cheng, Zhao Yang, Yong- Jun Lin, Zhong-Qiu Zhao, Yong-Li Jiang, De-Shuang Huang. Deep learning-based methods for person re-identification: a comprehensive review, *Neurocomputing*, 337 (14 April 2019), pp. 354–371.

[15]. Jean-Paul Ainam, Ke Qin, Guisong Liu, Guangchun Luo. Sparse label smoothing regularization for person re-identification, IEEE Access 7 (2019), pp. 27899–27910.

[16]. K. Zhou, Y. Yang, A. Cavallaro, T. Xiang. Omni-scale feature learning for person reidentification, ICCV, 2019.

[17]. Hong-Xing Yu, Ancong Wu, Wei-Shi Zheng. Cross-view asymmetric metric learning for unsupervised person re-identification, Proceedings of the IEEE International Conference on Computer Vision (2017), pp. 994–1002.

[18]. S. Xu, Y. C, K. Gu, Y. Yang, S. Chang, P. Zhou. Jointly attentive spatial-temporal pooling networks for video-based person re-identification, Proceedings of the IEEE International Conference on Computer Vision (2017), pp. 4733–4742.

[19]. C. Long, A. Haizhou, Z. Zijie, S. Chong. Real-time multiple people tracking with deeply learned candidate selection and person re-identification, ICME, 5, 2018, p. 8.

[20]. Yiheng Liu, Zhenxun Yuan, Wengang Zhou, Houqiang Li. Spatial and temporal mutual promotion for video-based person re-identification, arXiv preprint, 2018, arXiv:1812.10305.

[21]. Hong-Xing Yu, Ancong Wu, Wei-Shi Zheng. Unsupervised person re- identification by deep asymmetric metric embedding, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.

[22]. Z. Ji, S. Li, Y. Pang. Fusion-attention network for person search with free-form natural language, *Pattern Recogn. Lett.*, 116 (2018), pp. 205–211.

[23]. Cunyuan Gao, Rui Yao, Jiaqi Zhao, Yong Zhou, Fuyuan Hu, Leida Li. Structure-aware person search with self-attention and online instance aggregation matching, *Neurocomputing*,369 (2019), pp. 29–38.

[24]. Yan Lu, Zheran Hong, Bin Liu, Weihai Li, Yu. Nenghai, Dhff. Robust multi-scale person search by dynamic hierarchical feature fusion, 2019 IEEE International Conference on Image Processing (ICIP), IEEE 2019, pp. 3935–3939.

[25]. Angelique Loesch, Jaonary Rabarisoa, Romaric Audigier. End-To-End person search sequentially trained on aggregated dataset, 2019 IEEE International Conference on Image Processing (ICIP), IEEE 2019, pp. 4574–4578.

[26]. N. McLaughlin, J. M. del Rincon, and P. Miller, "Recurrent convolutional network for video-based person re-identi cation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 1325 1334.

[27]. I. Bouchrika. "A survey of using biometrics for smart visual surveillance: Gait recognition," in Surveillance in Action. Cham, Switzerland: Springer, 2018, pp. 3–23.

[28]. P. Zhang, Q.Wu, J. Xu, and J. Zhang. "Long-term person re-identication using true motion from videos," in Proc. IEEE Winter Conf. Appl. Com-put. Vis. (WACV), Mar. 2018, pp. 494–502.

[29]. Q. Leng, M. Ye, and Q. Tian. "A survey of open-world person reidentication," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.

[30]. F. M. Castro, M. J. Marín-Jiménez, N. Guil, and N. P. de la Blanca. "Automatic learning of gait signatures for people identication," in *Advances in Computational Intelligence (Lecture Notes in Computer Science)*, vol. 10306. Cham, Switzerland: Springer, 2017, pp. 257–270.

[31]. L. Yao, W. Kusakunniran, Q. Wu, J. Zhang, and Z. Tang. "Robust CNN-based gait verification and identification using skeleton gait energy image," in *Proc. Digit. Image Comput., Techn. Appl.* (DICTA), Dec. 2018, pp. 1–7.

[32]. D.Wu, S.-J. Zheng, X.-P. Zhang, C.-A. Yuan, F. Cheng,Y. Zhao,Y.-J. Lin, Z.-Q. Zhao, Y.-L. Jiang, and D.-S. Huang. "Deep learning-based methods for person re-identification: A comprehensive review," *Neurocomputing*, vol. 337, pp. 354–371, Apr. 2019.

[33]. Q. Leng, M. Ye, and Q. Tian. "A survey of open-world person reidentification," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.

[34]. A. Schumann and R. Stiefelhagen. "Person re-identication by deep learning attribute- complementary information," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, Jul. 2017, pp. 1435–1443.

[35]. Y. Lin, L. Zheng, Z. Zheng, Y.Wu, Z. Hu, C. Yan, and Y. Yang. "Improving person re- identification by attribute and identity learning," *Pattern Recognit.*, vol. 95, pp. 151–161, Nov. 2019.

[36]. S. Li, T. Xiao, H. Li, B. Zhou, D. Yue, and X. Wang. "Person search with natural language description," in Proc. 30th IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 5187–5196.