International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

Person Identification in Group Photographs with Artificial Intelligence: Survey

Miss. Priyanka G. Patil[1], Dr. Mrs. M. R. Dixit[2]

[1]Student, Dept. of Electronic Engineering KIT's collage of Engineering Kolhapur, Maharashtra, India.
[2]Professor, Dept. of Electronics & Telecommunication Engineering, KIT's collage of Engineering Kolhapur, Maharashtra, India.

Abstract - *Now a days, capturing photos using smartphone camera is very practical. Extremely when anybody hanging out with friends or family members, sometimes when they attending any occasions or functions and so on. At that time, they end up with taking many group photographs. But when browsing through all these group photographs, most of the times, a particular person is interested in only those photos in which he himself is present. While finding a person of interest can be vital for public safety, it is also a task of great visual complexity that requires continued attention, good identity or face detection and recognition skills and other required resources. The search target's photos were also real, being passport photos, custody images or social media images. By using this way we fulfil the gap between laboratory-based experiments and real-life CCTV search or images from real life. It signifies benefits for finding target identities in CCTV footage, which is a task that is conducted by the security officers around the world every day. So here this person identification with artificial intelligence is important.*

Identify and recognize particular person from group photographs with artificial intelligence.

Key Words: **Artificial Intelligence, Deep Learning, Neural Networks, Face Detection, Feature Extraction, Face Recognition, Image with Captioning**

1. INTRODUCTION

Person identification in photo or in CCTV footage is also important to recognize faces of celebrities, politicians, or other persons like VIPs in public spaces which is the most interesting work for reporters. In many countries, closed-circuit television (CCTV) surveillance is common in public spaces. The availability of CCTV footage has brought about significant changes policing area and finding the culprit or a gangster. The proposed work is to identify and recognize particular person in a group photographs by computing the features such as edge, line and center-surround features i.e. wavelet transform using the concept of deep neural networks in artificial intelligence. For this process collection of image data is required. From Image acquisition face detection is done. By using feature extraction we classifying the image. KNN classifier method will be using for classify the particular image from other images. From classification we will recognize the image & proceed to image enhancement. Then Image will display with the captioning.

2. LITERATURE

The method implemented a Homogeneous Augmented Tri-Modal (HAT) learning method for VI-ReID, where an auxiliary grayscale modality is generated from their homogeneous visible images, without any additional training process. Especially, they have solved the tri-modal feature which is learning from both multi-modal classification and multi-view retrieval perspectives. Incorporated with two invariant regularizers, HAT concurrently reduces multiple modality variations. Furthermore, author have developed a weighted tri-directional ranking loss to improve the relative distance across the cross-modality positive and negative triplets [1].

System designed to improve the robustness of Metric Learning by Accelerated Proximal Gradient (MLAPG), a person re-ID algorithm, called Asymmetric-MLAPG which is built on the basis of asymmetric metric. Usually traditional metric learning ignores the taking environment of the person images. Where an asymmetric metric learning aims to find the feature matrix of each camera and map the vectors of images to a common area space. Author have rebuilt MLAPG model by Asymmetric metric. For adding a regularization term which controls the influence of inconsistency of metric, an improved MLAPG algorithm based on asymmetric metrics is verified. An asymmetric metric introduced to make each camera learn its own feature mapping matrix of images which improves the robustness of the algorithm. Where it will increase the number of iterations [2].

Analysis of a networked video surveillance system, where the videos collected from different surveillance cameras are stored in a centralized server. Real time analysis & verification of these surveillance videos can be helpful for preventing crimes in the areas under surveillance. Authors have discussed a method for person re-identification from the videos collected from surveillance cameras. Here the

main task is that process had face detection from videos and prediction of persons using Convolutional Neural Network models developed using the cropped face images from the face detection stage. The main contribution of their work is the comparison of two approaches for face detection namely Viola Jones algorithm and Multitasking Convolutional Neural Networks [3].

The pedestrian alignment network (PAN) which allows discriminative embedding learning pedestrian alignment without extra annotations. Authors have observed that when the convolutional neural network learned about the discriminate between different identities, the learned feature mapped usually exhibit strong activations on the human body rather than the background. Designed network system takes benefit of this attention mechanism to adaptively locate and adjust pedestrians within a bounding box [4].

Implementation of a novel, simple and efficient person re-id network is called MPLFN. This network is a combination of two tasks that is classification task and metric learning task. The classification task involves uniformly partition of N feature parts from an image and evaluate the person classification loss in each part separately. Metric learning task involves recalculation of distance of two images by the shortest path between two sets of feature parts. Then the distances are put into a triplet loss to perform a dynamic part alignment during the training. With the joint learning of these two tasks, the performance of the network is significantly magnified [5].

Analysis of the person re-identification into a regression process and found probability that person in the images belong to the same identity. First, they have implemented a deep regression model, named deep regression neural network integrating adaptive multi-attribute fusion method (DRNN-AMAF), it can made the person re-identification as regression analysis. And another attributes are taken as the basis of this model for finding the probability of persons belonging to the same identity. Multi-label fusion method which is based on the idea of Bayesian inference, which made the attribute labels suitable for regression tasks [6].

Implementation of an Android platform based photo grouping application named “Euphoria Grouping” (EUG) using Neural Networks. They have automated the process by planning and generating an Android application named “Euphoria Grouping” (EUG). Deep learning uses neural networks. There are different types of neural networks such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), etc. EUG application automated the process of detecting faces and identified persons from a group photo using Deep Learning. Two different CNN models, viz, Custom Built and OpenFace CNN are used for person identification. From the results, they concluded that the performance and the accuracy of the OF-CNN model is 93.67% which is higher than CB-CNN model [7].

Creation of a new deep hashing network called Deep Semantic Structured Hashing (Deep SSH) is to design the semantic structured representation of human. In the designed Deep SSH framework, both the mid-level human attributes and the high-level ID labels are used. Based on those semantic structured hash code and the attribute labels, they have develop a decoder which is used to find the partial hash code corresponding to the specified attributes. The Deep SSH method designed a part based structure which can utilize the semantic information better. Meanwhile, they have made the full use of the person attribute datasets to assemble the structured hash codes embedding the attribute information [8].

Implementation of a body structure based triplet convolutional neural network (BSTCNN) for person re-identification. To acquire both feature representation and similarity metric concurrently, deep metric learning methods using triplet convolutional neural network have been referred in person re-identification. Especially, a 4-branch CNN architecture is made to understand the features from different parts of body. They have implemented an improved triplet convolution neural network for person re-identification. Their network contains two novel elements: a weighted distance layer that acquire weights for each body part, an improved loss function that speeds up convergence and increases the performance [9].

Analysis of a person re-identification, most of which however it neglect the matching efficiency. Despite their efficiency, these works ignore cross-camera variations, which severely deteriorate the final matching accuracy. To direct the above issues, author found a novel hashing based method for fast person re-identification, namely Cross-camera Semantic Binary Transformation (CSBT). CSBT transforms the original high-dimensional feature vectors into compact identity preserving binary codes. CSBT first recruit a subspace projection to mitigate cross-camera variations, by maximizing intra-person similarities and inter-person discrepancies. Binary coding scheme is introduced via seamlessly incorporating both the semantic pairwise relationships and local affinity information [10].

Analysis of an efficient face identification technique in a group photograph. They used SURF i.e., Speeded Up Robust Features technique which is used to identify a person in a group photograph. SURF is contain of a feature detector based on a Gaussian second derivative mask, and a feature descriptor that relies on local Haar wavelet responses. The main concept of these features is to find nearest neighbors for a numerical vector. The SURF detector finds its attention on blob-like structures in the image. These structures can be observe at corners of objects, but also at locations where light speckles are present. To relate each

feature, SURF sum up the pixel information within a local neighborhood [11].

Implementation of an open source face recognition system named XFace for Android mobile phone platform for identification. It gives a feasible solution to meet these challenges. The designed face recognition approach includes face detection, eye detection, pre-processing for ROI (Region of Interest), LBP (Local Binary Pattern) feature extraction, feature dimensionality reduction based on PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis), and minimum distance classifier. The implementation of the proposed system is mainly based on Open CV (Open Source Computer Vision) SDK [12].

Implementation of a novel approach of learning mid-level filters from automatically discovered patch clusters for person re-identification. The effectiveness of middle level filter is verified on the VIPeR dataset and the CUHK01 dataset. The mid-level features are complex to existing handcrafted low level features, and enhanced the best Rank-1 matching rate on the VIPeR dataset by 14%. They have explored different discriminative abilities of local patches by introducing pAUC score. Coherent patch clusters are created by pruning hierarchical clustering trees, and a simple but effective cross-view training strategy is introduced to learn filters that are view invariant and discriminative in distinguishing identities [13].

A novel filter pairing neural network (FPNN) to handle misalignment, geometric transforms, occlusions and background clutter. In practice to existing works, which uses handcrafted features, the method automatically captures the features optimal for the re-identification task from data. Its deep architecture makes it possible to model a mixture of complex photometric and geometric transforms. They have formed a largest benchmark re-id dataset with 13,164 images of 1,360 pedestrians. The neural network significantly outperforms state-of-the-art methods on this dataset. Some effective training strategies are found to train the network well [14].

Analysis of an interactive algorithm to automatically segment out and recognize a person's face from a group photograph. Author have designed face recognition systems to recognize individuals from the group photograph containing multiple people. It is related to as face location, face extraction, or face segmentation. So, they have introduce this method. The method have a fast, reliable, and effective algorithm that creates the spatial distribution characteristics of human skin color [15].

Designing of a system for detection faces in clumsier group photos. In this system the basic image processing concepts such as color segmentation, image segmentation and template matching techniques will be used in several stages. After face detection, the faces should be recognized. The eigen subspaces of the face image is analysed. From this the features are extracted. The features of the detected face will be sum up with the features of the existing database faces detected and recognized. The analysed system having two modules: face detection and face recognition. In a group photograph the face regions are detected. Then these detected face is recognized by sum up with the database face image [16].

Analysis of the person re-identification under single shot, and implemented a new method on the basis of metric methods and classifiers, cutting pedestrians into parts and performing local metric learning on each subset of component, meanwhile, carrying out global metric learning combined or both with SVM classifier in the entire training set. Person re-identification mentions to pedestrian matching under the surveillance of non-overlapping multi-camera, which is usually used in criminal investigation work and image retrieval. However, the local metric learning method usually needs to complete complex convex optimization problems, which is not come along with practical application [17].

Implementation of language modelling and nearest neighbour approaches to context-based person identification, in addition to novel face color and image color content-based features. Identification of people in photographs is an important need for users of photo management systems. They have designed MediAssist, one such system which facilitates browsing, searching and semi-automatic annotation of personal photos, using analysis of both image content and the context in which the photo is captured [18].

Creation of a novel and efficient facial image representation based on local binary pattern (LBP) texture features. The face image is classified into several regions from which the LBP feature distributions are extracted into an enhanced feature vector to be used as a face descriptor. A novel and efficient facial representation is designed. It is based on dividing a facial image into small sections and developing a description of each region using local binary patterns. The LBP operator has been very useful in different applications such as texture classification, image retrieval, etc. [19].

Analysis of a new human skin color model in YCbCr color space and its application to human face detection. Skin colors are modeled by a set of three Gaussian clusters, each of which is described by a centroid and a covariance matrix. The centroids and covariance matrices are derived from large set of training samples after a k-means clustering process. They have concluded that presented a skin color-based face detection algorithm that displays a human skin color model, which takes into account the luminance Y in classifying skin and non-skin pixels. In their approach, the distribution of human skin colors in YCbCr

space is modeled with 3 Gaussian clusters that approximately equals to 3 levels of luminance: low, medium and high [20].

3. CONCLUSION

Person identification in group photos by using deep learning neural network in artificial intelligence. And after face detection, person will be recognized and then image is display with captioning.

ACKNOWLEDGEMENT

I would like to thank Dr. Mrs. M. R. Dixit for the elegant suggestion, guidance and support.

REFERENCES

Mang Ye, Jianbing Shen, Senior Member, IEEE and Ling Shao, "Visible-Infrared Person Re-Identification via Homogeneous Augmented Tri-Modal Learning", IEEE Transactions on Information Forensics and Security, June 2020.

[1] Wenbi Rao, Menghan Xu, Junwei Zhou, "Improved Metric Learning Algorithm for Person Re-Identification Based on Asymmetric Metric", 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA).

[2] Vimala Mathew, Tom Toby, Anu Chacko, Udhayakumar A, "Person re-identification through face detection from videos using Deep Learning", IEEE, 2020.

[3] Zhedong Zheng, Liang Zheng, Yi Yang, "Pedestrian Alignment Network for Large-scale Person Re-Identification", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 29, No. 10, October 2019.

[4] Zimian Wei, Wenjing Yang, Wanrong Huang, Huadong Dai and Dongsheng Li, "Person Re-Identification Based on Multi-Parts of Local Feature Network", IEEE, Volume 7, 2019.

[5] Yingchun Guo, (Member, IEEE), Kunpeng Zhao, Xiaoke Hao, (Member, IEEE), and Ming Yu, (Member, IEEE), "Deep Regression Neural Network for End-to-End Person Re-Identification", IEEE, Volume 7, 2019.

[6] Prof. Avani Sakhapara, Prof. Dipti Pawade, Mr. Saumil Dedhia, Ms. Twinkle Bhanushali, Mr. Vinit Doshi, "Machine Learning Based Approach for Person Identification in Group Photos", Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018.

[7] Ya Zhao, Sihui Luo, Yezhou Yang, Mingli Song, "Deepssh: Deep Semantic Structured Hashing For Explainable Person Re-Identification", Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies(AZFT) Zhejiang University, Hangzhou 310027, China, ICIP 2018.

[8] Hong Liu, Weipeng Huang, "Body Structure Based Triplet Convolutional Neural Network for Person Re-Identification", Key Laboratory of Machine Perception, Shenzhen Graduate School, Peking University, ICASSP 2017.

[9] Jiaxin Chen, Yunhong Wang, Jie Qin, Li Liu, Ling Shao, "Fast Person Re-identification via Cross-camera Semantic Binary Transformation", IEEE Conference on Computer Vision and Pattern Recognition, IEEE 2017.

[10] Jyoti Dalal, Mahendra Singh Meena, Prof. Priti Singh, "Person Identification In A Group Photograph Using Surf Features", International Journal of Innovations & Advancement in Computer Science IJIACS, ISSN 2347 – 8616 Volume 4, Issue5 May 2015.

[11] Jiawei Hu, Liangrui Peng, Li Zheng, "XFace: A Face Recognition System for Android Mobile Phones", IEEE 3rd International Conference on Cyber-Physical Systems, Networks, and Applications, IEEE 2015.

[12] Rui Zhao, Wanli Ouyang, Xiaogang Wang, "Learning Mid-level Filters for Person Re-identification", IEEE Conference on Computer Vision and Pattern Recognition, IEEE 2014.

[13] Wei Li, Rui Zhao, Tong Xiao, Xiaogang Wang, "DeepReID: Deep Filter Pairing Neural Network for Person Re-Identification", IEEE Conference on Computer Vision and Pattern Recognition, IEEE 2014.

[14] Kavita Shelke, "Face Recognition from Group Photograph", International Journal of Engineering and Innovative Technology (IJEIT) Volume 3, Issue 1, July 2013.

[15] Robert Singh, A. Suganya, "Efficient Tool for Face Detection and Face Recognition in Color Group Photos", VIT University, Vellore, India, 2011.

[16] Xiaofeng Gu, Xiexin Zou, Jia Liu, Lixue Zhang, "Person Re-Identification By Using A Method Combining DPM and SVM", International Centre for Wavelet Analysis and Its Applications, School of Information and Software Engineering, University of Electronic Science and Technology of China, 2011.

[17] Neil O'Hare and Alan F. Smeaton, Member, IEEE, "Context-Aware Person Identification in Personal Photo Collections", IEEE Transactions on Multimedia, Vol. 11, No. 2, February 2009.

[18] Timo Ahonen, Student Member, IEEE, Abdenour Hadid, and Matti Pietika¨ inen, Senior Member, IEEE, "Face Description with Local Binary Patterns: Application to Face Recognition", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 28, No. 12, December 2006.

[19] Son Lam Phung, Abdesselam Bouzerdoum, Douglas Chai, "A Novel Skin Color Model In Ycbcr Color Space And Its Application To Human Face Detection", Visual Information Processing Research Group Edith Cowan University, Westem Australia, IEEE KIP 2002.