International Current Journal of Engineering and Science (ICJES)
Volume 3, Issue 10, pp 19-24, October-2024
DOI: https://doi.org/10.47001/ICJES/2024.310005
ISSN (Online): 2583-7257

Person Identification from Surveillance Camera Images Using Three-Layer Convolutional Neural Network

[1]Thisara N.D.N.A, [2]Jayasuriya D.G.T, [3]Kumara V.A.S.M, [4]Sandeepa H.D.S.R, [5]Muthukudaarachchi A.H.M

[1,2,3,4,5]Department of Information Technology, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

***Abstract* - Deep learning represents an important field of machine learning, most notably the use of three-layer neural networks. These networks aim to mimic how the human brain works; however, they still fail to effectively "learn" from broad datasets. Advanced image surveillance personal re-identification (ReID) allows multiple cameras to identify the same person. This process is complicated by factors such as disorientation, unique camera perspectives and differences in human posture. The challenges faced by human-identification (ReID) are important, especially to address these issues due to unrestricted spatial mismatches between two images due to changes in perspective and pedestrian positions, and labeled noise from clustering methods, convolutional Neural networks (CNNs), using a preprocessing method based on reinforcement learning, use local pair wise internal representation interactions to organize a specific task sequence over at the point corresponding to two images This method is considered to be the best method for human ReID and is performed in accordance with the most effective features. In addition, it is important to provide examples of widely used datasets, explore the strengths and weaknesses of different approaches, and compare the performance of specific algorithms on newly obtained image datasets that can then be imaged with a CNN has been used to train deep learning models for face recognition. CNNs are particularly useful for computer vision (CV) processing, image recognition, and image classification, as they provide highly accurate results, especially when processing large amounts of data. Compared with the existing methods, the proposed method achieves accuracy rates of 96.0% and 89.0%, respectively.**

Keywords: Person Re-Identification, Deep-Metric Learning, Local Feature-Learning, Generative Adversarial Learning, Sequence Feature-Learning.

I. INTRODUCTION

Person recognition (re-ID) is a complex computer vision task with the goal of identifying persons from multiple images or video images captured by non-overlapping or continuous cameras Interested in what it can of use in personal tracking, public safety, and video surveillance has increased recently. Person recognition aims to smoothly and consistently match a person's query image with corresponding images in an extensive gallery collection, even when there are noticeable changes in appearance, location, lighting, and in the prevention Common features are excluded from low-level visual information such as color, texture, and appearance and training discrimination systems to match and rank individuals.

However, despite these important advances, human identification remains a challenging task, especially given the challenges of managing large-scale variables and collecting fine-grained features under mundane conditions in itself because. Deep neural networks have been used with great success in computer vision tasks such as object recognition and image segmentation. Using deep learning algorithms, researchers have surpassed the ability of traditional methods to achieve state-of-the-art performance in the form of self-recognition

1.1 Person re-identification

In the computer vision task of identity re-identification (re-ID), subjects are identified and compared between multiple images or video images captured by uninterrupted or overlapping cameras In recent years, there has been considerable interest in its applications in areas as diverse as public safety, human tracking and video surveillance. Human re-ID attempts to accurately and quickly identify a subject from a query image, compared to similar images in a gallery, where apparently different form, attitude, illumination, and occlusion Human recognition presents different challenges than other computer vision tasks there are performances, such as changing clothes, camera views and environmental conditions. Real-world environments also introduce additional challenges to the project since there may be a number of images in the gallery with poor labeling or overlapping prior, human re-ID methods of artificially generated feature representations to match and rank persons, such as color histograms, texture descriptors or geometric attributes, but these methods were often unable to extract the subtle and subtle information needed to accurately re-identify individuals.

International Current Journal of Engineering and Science (ICJES)
Volume 3, Issue 10, pp 19-24, October-2024
DOI: https://doi.org/10.47001/ICJES/2024.310005
ISSN (Online): 2583-7257

II. LITERATURE SURVEY

H. Wanglee [1], et al. As shown in this study, computer vision, pattern recognition, signal processing, embedded computing, communication, and image sensors are all involved in various tasks involving multi-camera video monitoring so intelligent This review reviews recent advances in related technologies in computer vision and pattern recognition. Topics covered include multi-camera analysis, object recognition, collaborative video viewing with automatic camera, multi-camera measurement, inter-camera network topology computation and faced with technical challenges in providing an in-depth review of the comparison of various solutions. It focuses on the relationships and interactions between modules in different contexts and application scenarios. According to recent research, a number of problems can be solved peacefully to improve accuracy and performance. Rapid surveillance technologies increase the size and complexity of surveillance cameras, as well as the management of crowded and complex systems this article focuses on solving this new problem so. Intelligent video surveillance has become one of the most evolving areas of computer vision. The objective is to identify, track and identify objects of interest in the large number of images captured by security cameras, and to understand and analyze their movements. Video surveillance can be useful in both public and private contexts many in monitoring scenes from roads, train stations, parking lots, stores, shopping malls, and workplaces—both inside and outside—is essential for these applications.

In this document, L. JiYang, [2] et al. Recent advancements in computer vision have led to an increased focus on person re-identification systems, commonly referred to as person Re-ID. These systems play a vital role in intelligent visual surveillance, with applications that significantly contribute to public safety. Person Re-ID technology enables the identification of individuals across a network of cameras, even when the cameras do not overlap, which is particularly challenging due to variations in camera angles. Factors such as lighting changes, occlusion, and positional differences further complicate this task. To tackle the challenges associated with person Re-ID, various methods for creating handcrafted features have been developed. In recent years, deep learning prompting numerous studies to explore its application in enhancing person Re-ID performance. This study aims to provide a comprehensive overview of recent research advocating for the integration of deep learning techniques in person Re-ID systems. It also discusses the public datasets utilized for evaluating these systems. The study concludes by highlighting ongoing challenges and suggesting future directions for the advancement of person Re-ID technologies. The growing importance of effective video surveillance for security purposes, including forensic investigations and crime prevention, has spurred significant governmental efforts to enhance surveillance technologies. Automated video monitoring and analysis remain critical components of intelligent video surveillance systems, although they require considerable time and effort from human operators.

Person re-identification serves as a vital component of advanced video surveillance systems. It is characterized as the mechanism through which a network of non-overlapping cameras, situated in diverse geographic areas, can identify and recognize the same individual. The challenge of Person Re-ID arises from the fact that footage is captured by cameras that do not overlap, often under varying conditions. Consequently, relying solely on primary biometric data, such as facial recognition, proves to be inadequate. While research predominantly emphasizes an individual's appearance, it also encounters considerable visual ambiguity stemming from variations both within and across different classes. Despite being one of the most significant tasks in intelligent video surveillance, Person Re-ID remains a complex endeavor. This survey examines Person Re-ID systems that utilize deep learning techniques, illustrating the development of a generic architecture applicable to both deep learning and traditional learning systems. Numerous recent studies have adopted deep learning approaches to address the shortcomings of manual methods.

L. Sheng, W. Wang [3] and others. Person re-identification, commonly referred to as re-ID, has been integrated into this system and is gaining increasing popularity within the research community due to its significant applications and benefits. The focus of re-ID is to identify a specific individual across different camera feeds. Initial studies primarily emphasized handcrafted algorithms and small-scale evaluations. However, recent advancements have led to the creation of large-scale datasets and the application of deep learning algorithms that utilize extensive data resources. In light of various tasks, we classify the majority of existing re-ID methodologies into two categories: image-based and video-based, with an examination of both handcrafted and deep learning approaches for each category. Additionally, this paper discusses two innovative re-identification tasks that are particularly pertinent to real-world applications: end-to-end re-ID and rapid re-ID within extensive galleries. It provides an overview of the evolution of person re-ID and its connections to image classification and instance retrieval, highlights key future research directions in end-to-end re-ID and swift retrieval in large galleries, and briefly addresses several significant yet underexplored challenges. Furthermore, it

International Current Journal of Engineering and Science (ICJES)
Volume 3, Issue 10, pp 19-24, October-2024
DOI: https://doi.org/10.47001/ICJES/2024.310005
ISSN (Online): 2583-7257

encompasses a wide range of both manually crafted systems and advanced techniques for image- and video-based re-identification. In the narrative recounted by Homer, Menelaus embarked on a journey to appease the gods and ensure a safe return home, yet he encountered difficulties in navigating his way back after the Trojan War. His objective was to capture Proteus and compel him to disclose the means of returning home. Menelaus successfully apprehended Proteus while he was resting among the seals after emerging from the sea, despite Proteus's attempts to evade capture by transforming into various forms, including water, trees, lions, snakes, and leopards. Ultimately, Proteus was forced to provide truthful answers. This tale may represent one of the earliest instances of an individual reclaiming their identity despite undergoing significant physical changes.

In 1961, Alvin Plantinga contributed significantly to the discourse on the connection between mental states and behavior by offering one of the initial definitions of re-identification. He articulated that "to re-identify a particular is to recognize it as (numerically) the same entity as one previously encountered." Consequently, extensive research and documentation regarding person re-identification have been conducted across various fields, including logic, psychology, and metaphysics. These investigations are fundamentally rooted in Leibniz's Law, which posits that "there cannot be distinct objects or entities that share all their properties." The role of person re-identification in contemporary computer vision parallels that of earlier studies.

Using this method has been suggested by J. Redshimon and A. Farhuadi. [4] have proposed the utilization of this method. We are in the process of implementing several enhancements for YOLO. Numerous minor design modifications have been made to improve its functionality. Additionally, we have trained a new, highly efficient network. Although this version is slightly larger, it offers greater accuracy. However, the increase in size is not substantial, so there is no cause for concern. At times, one may find themselves improvising for an entire year. My research efforts this year were limited; I logged into Twitter on several occasions and conducted some experiments with GANs. I was able to carry forward some momentum from the previous year, which facilitated improvements to YOLO. Nonetheless, the changes are relatively unremarkable, consisting of a few minor adjustments aimed at enhancement. I have also provided some assistance to a few individuals with their academic pursuits, which is a primary reason for our current position. As we approach our deadline for camera readiness, it is essential to reference a couple of my informal modifications to YOLO, despite lacking a formal source. Prepare for a technological report! Technical reports are advantageous as they do not require an introduction, given that the audience is already aware of the subject matter. Thus, the conclusion of this introduction will serve as a guide for the remainder of the document. We will commence by detailing the context of YOLOv3, followed by an assessment of our performance. Additionally, we will share some of our unsuccessful attempts. Finally, we will discuss the significance of our findings. Concatenation is employed to integrate our upsampled features with a preceding network feature map, enabling the extraction of more substantial semantic information from the upsampled features alongside finer-grained details from the earlier feature map. Prepare for a technological report.

The primary advantage of technical reports is their inherent clarity regarding purpose, which eliminates the need for an introductory section. Thus, the conclusion of this introductory segment will serve as a guide for the subsequent content. We will commence by detailing the context of YOLOv3, followed by an assessment of our performance. Furthermore, we will discuss some of the less successful outcomes. To analyze the combined feature map, we will incorporate additional convolutional layers. Ultimately, we expect to generate a tensor that is twice the size yet remains comparable. This design will be replicated to predict bounding boxes for the final scale. As a result, our predictions for the third scale will leverage all prior computations along with the intricate details from the earlier stages of the network. Each bounding box will utilize multilabel classification to forecast the potential classes contained within. Our findings indicate that satisfactory performance can be achieved without employing a softmax function; therefore, we utilize independent logistic classifiers. During the training phase, we apply binary cross-entropy loss for the class predictions.

In this study, W. Luang, Angualov et. al. [5] have suggested. We introduce a novel approach to image-based object detection utilizing a single deep neural network. This method, designated as SSD, the network amalgamates predictions from multiple feature maps of differing resolutions, enabling it to effectively manage objects of varying sizes. SSD simplifies the detection process compared to methods that rely on object proposals, as it consolidates all computations experimental evaluations on the COCO, ILSVRC, and PASCAL VOC datasets demonstrate that SSD offers a cohesive framework for both inference and training, significantly outperforming methods that necessitate an additional object proposal stage in terms of speed, while also achieving competitive accuracy. Variants of this methodology are employed by the most advanced object detection algorithms available today, which utilize high-quality classifiers following the resampling of features or pixels for each bounding box, based on specific assumptions regarding bounding boxes. This pipeline has

International Current Journal of Engineering and Science (ICJES)
Volume 3, Issue 10, pp 19-24, October-2024
DOI: https://doi.org/10.47001/ICJES/2024.310005
ISSN (Online): 2583-7257

consistently excelled in detection benchmarks, beginning with the Selective Search methodology and extending to the latest leading results on PASCAL VOC, COCO, and ILSVRC detection, all of which are based on Faster R-CNN but incorporate richer features.

III. PROPOSED SYSTEM

By using this method, the CNN extracts global feature representations from the input photos that represent the general characteristics of a person's appearance. The ability to match and compare individuals across multiple photos is made possible by these global traits, and this is essential for person ReID. The method also employs a mechanism called learned alignment regions, which recognises specific regions in the images that are relevant to the individual's ReID. CNN is able to focus on specific topics and pinpoint characteristics that are particular to them. This improves the accuracy of the ReID process by emphasising the salient features of the subject's appearance. To facilitate a better understanding of sequential spatial correspondences between picture pairs, a location network is presented.

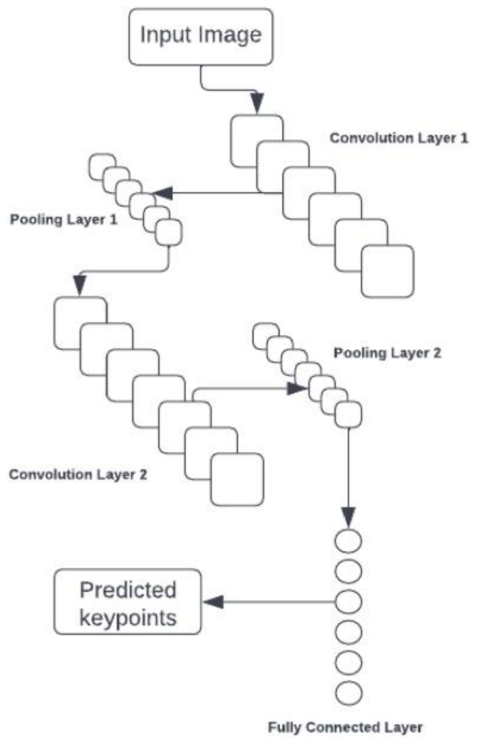

Figure 1: CNN face recognition algorithm

Although these advanced technologies exhibit remarkable accuracy, they have been found to be excessively demanding in terms of computational resources for embedded systems and inadequate in speed for real-time applications, even when employing the latest advancements. The most rapid Faster R-CNN, achieves only a limited number of frames per second (FPS), while the detection rates of these techniques are frequently assessed in seconds per frame (SPF). We have made several attempts to address this issue. Ultimately, we will discuss the significance of all components involved. Additionally, we integrate our upsampled features with a prior network feature map through concatenation.

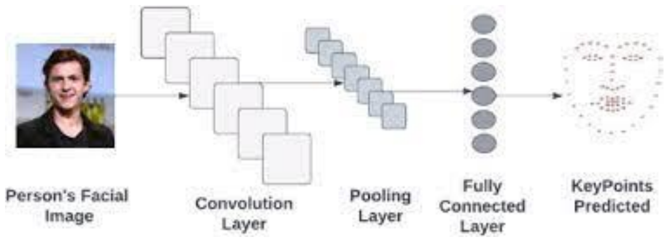

Figure 2: Person facial recognition

This network can create spatial correspondences between different picture pairs by learning to make sequential judgements because it is based on reinforcement learning. By doing this, the location network is able to align and match the features that are retrieved from the images. The learned alignment areas and CNN features are then fed into a Deep Learning Algorithm (DLA). To be used in future tasks such as person matching, retrieval, or tracking, these attributes need to be processed and organised by the DLA.

3.1 Advantages of Proposed System

Global Feature Representation: The system extracts global feature representations from input photos using a Convolutional Neural Network (CNN) to capture general aspects of an individual's appearance. This enables a thorough comprehension of the general appearance, which facilitates matching and comparing people in various pictures.

Learned Alignment Regions: By emphasising prominent aspects of the subject's appearance, the system uses a mechanism to identify particular regions in the images that are pertinent to each individual ReID, improving accuracy. This focused strategy enhances the precision of person identification and matching.

Focused Feature Extraction: The accuracy of the ReID procedure is improved by the CNN's capacity to concentrate on particular subjects and identify traits unique to each person. Through prioritising pertinent features, the system lowers noise and enhances matching quality. The utilisation of a Location Network for Spatial Correspondences can aid in the comprehension of sequential spatial correspondences between picture pairs. Reward learning teaches the network to make sequential decisions, which improves matching accuracy by aligning and matching features retrieved from images.

International Current Journal of Engineering and Science (ICJES)

Volume 3, Issue 10, pp 19-24, October-2024

DOI: https://doi.org/10.47001/ICJES/2024.310005

ISSN (Online): 2583-7257

Integration with Deep Learning Algorithm (DLA): By incorporating the CNN features and learned alignment areas into a Deep Learning Algorithm (DLA), additional attribute organisation and processing are made possible for tasks like tracking, retrieval, and person matching.

IV. FUTURE WORK

Future work on ordered or order-less person re-identification using CNN algorithms may focus on a variety of subjects. Initially, research may be done on advanced network topologies and training techniques that effectively capture and represent temporal relationships for re-identification of ordered individuals. This entails investigating transformer-based models, recurrent neural networks, and attention processes to better use temporal information. Secondly, if robust techniques are created to handle challenging scenarios such as occlusions, shifting viewpoints, or variations in illumination, both ordered and orderless systems might function better. The application of domain adaptation or transfer learning techniques is another important way to improve generalisation across several datasets or domains.

V. CONCLUSION

It was proposed that a Convolutional Neural Network (CNN) be used to solve the person re-identification problem. The constructed CNN architecture has one convolution layer, which is contrasted with additional convolution layers. Consequently, information about picture attributes might be included in the feature representations of our model. The architecture is trained using a set of generate features with the aim of bringing examples of the same person closer to a single camera while using structured samples to keep examples of different people farther apart in the learnt feature space. On benchmark datasets, this model mostly performed well. It calculates more precisely and quickly. In order to handle additional tasks in the future, such as picture and video retrieval, we intend to broaden our framework and methodology. The developed method shows that the characteristic's features are important cues to the person's re-id work, and their auxiliary data could enhance the pedestrian's ability to describe themselves.

REFERENCES

[1] H. Wanglee, Recognition of Pattern letters 40 (2020) 7–18, A review on video surveillance using intelligent multiple camera.

[2] L. JiYang, X. Fuhan, S. Zuwang, H. WuLing, Re-identication of individual persona by examination using deep learning techniques, Act Automatica Scientifica 35 (2019) 2022–2045.

[3] Re-identification of person in Past, present and future, L. Sheng, W. Wang, and Y. G. Haptumann, 2020, arXiv preprint arXiv:1611.02764.

[4] J. Redshimon and A. Farhuadi, An enhancement of objects in image using incremental spanning techniques, arXiv preprint arXiv:1406.02722, 2021

[5] Multibox Single shot detector MSSD, W. Luang, E. Angualov, D. Erikuan, C. Shengedy, S. Ruan and Fu Berg, Proceedings of the European Conference on Computer Vision, 2021, pp. 20–35.

[6] Research on person re-identification based on supervised mapping by L. LianQi, P. Yuwang, and Y. Guao, Journal of Software, 10 (2020), 2867–240

[7] Zhuang, Zhing, Dao, S., D., and Lee, Proceedings of the IEEE Conference on Pattern Recognition and Computer Vision, 2021, pp. 3352–3361, Re-ranking using k-reciprocal encoding scheme for person re-identification.

[8] Salience matching technique for Person re-identification by: R. Zaong, W. Oiuying, and X. Wung, Proceedings of the IEEE International Conference on Computer Vision, 2021, pp. 2519–2528.

[9] Novel human re-identification based on pyramid color feature topology, H. Hiu, W. Fang, G. Zeng, Z. Hu, B. Li, Multimedia Navigation Tools, 76 (2020) 26633–26646.

[10] N. Ahuja, T. Huang, M. Dikmen, E. Akbas, and Pedestrian identification using an acquired metric, European Conference on Computer Vision and New Technologies Proceedings, 2021, pp. 501-412.

Citation of this Article:

Thisara N.D.N.A, Jayasuriya D.G.T, Kumara V.A.S.M, Sandeepa H.D.S.R, & Muthukudaarachchi A.H.M. (2024). Person Identification from Surveillance Camera Images Using Three-Layer Convolutional Neural Network. *International Current Journal of Engineering and Science - ICJES*, 3(10), 19-24. DOI: https://doi.org/10.47001/ICJES/2024.310005

****