A Review of Latest Trends and Technologies in the Field of Facial Re-recognition in Surveillance Cameras

Ait Sadi Said Ameziane

ABSTRACT
Deep Learning-driven advances in facial re-identification are transforming computer vision. The increasing need for precise identity identification in surveillance, law enforcement, and public safety is being met by this advancement. Complex face feature extraction is automated by Deep Learning, particularly with Convolutional Neural Networks (CNNs), improving recognition under difficult circumstances.

This paper focuses on the most recent developments and the influence of Deep Learning on facial re-identification. For scholars, practitioners, politicians, and industry actors, it's a vital resource that emphasizes the need of staying current for the best possible system use. To sum up, the combination of Deep Learning with facial re-identification offers accurate, dependable, and effective identity recognition. It is essential for security, law enforcement, and public safety that technology advances.

Keywords: Deep Learning, facial re-identification, computer vision, surveillance, law enforcement, public safety, CNNs, identity recognition, recent developments

1. INTRODUCTION
1.1: Introduction to Deep Learning
A technological advancement in computer science can be seen in the terms Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL). According to Ongsulee (2017), the term "AI" refers to the broad idea of creating computers or systems capable of intelligent behavior, which covers a variety of methodologies and approaches. The goal of machine learning (ML), a subset of artificial intelligence, is to train algorithms to learn from data and develop over time. On the other hand, deep learning is a branch of machine learning that focuses on using deep neural networks, particularly deep artificial neural networks that are modeled after the structure of the human brain [2]. These networks are perfect for tasks like image and speech recognition because they consist of a large number of interconnected layers of nodes (neurons), which enable the automatic extraction of characteristics and patterns from data [3]. Figure 1.1 depicts the distinctions between AI, Machine Learning, and Deep Learning.
1.2: Use of Deep Learning in Facial Recognition

Deep Learning has transformed facial recognition by allowing it to automatically learn and extract complex facial traits and patterns from images or videos. This outperforms traditional methods that rely on manual feature engineering. Convolutional Neural Networks (CNNs), a sort of deep neural network, have been instrumental in this shift. CNNs can interpret visual input quickly, making them excellent for facial analysis jobs [4]. The application of deep learning has dramatically increased the accuracy and durability of facial recognition systems, allowing them to succeed even in difficult circumstances with varying illumination, position, and occlusions [5].

The below fig 1.2 shows the flow how CNN can be used for facial recognition. By using multiple convolution layers, this is achieved.

![Fig 1.2: Facial Recognition flow using CNN (Source: Shinde and Shah, 2018)](image-url)
1.3: Facial Re-identification and use of Deep Learning in Facial Recognition

Facial re-identification is a subfield of computer vision that focuses on recognizing and verifying individuals in a variety of camera angles, environmental conditions, and contexts. Several compelling grounds support its significance [6].

To begin, in the field of surveillance, where the number of operational cameras surpassed 770 million in 2021 and is expected to hit 1 billion by 2023, there is an urgent need for effective methods to follow individuals smoothly across a plethora of video feeds. Conventional methodologies frequently struggle to deal with the complexities of real-world settings [7].

Second, it is impossible to emphasize the importance of security and crime prevention. According to global data, there were around 3.5 million burglaries in 2020, with approximately 4.5 million violent crimes documented in the United States alone [9]. Facial re-identification aids law enforcement authorities in the resolution of criminal cases and the enhancement of overall security measures.

Furthermore, the deployment of facial re-identification technology in congested public venues such as airports and stadiums is critical for public safety [9]. It enables rapid threat detection and mitigation, which is especially important in light of escalating security concerns [10].

Deep Learning, as a transformative technology, is critical in addressing the issues associated with facial re-identification. It excels in dealing with the inherent variety in facial appearances caused by factors such as lighting, viewing angles, and facial expressions [11, 12]. Furthermore, its ability to successfully scale for processing the massive volumes of visual data provided by surveillance cameras is critical for automating and properly identifying individuals. Also, Deep Learning models are adaptable and have strong generalization capabilities, ensuring consistent performance even in complex and tough surveillance contexts [3, 4].

1.4: Significance of the work

This review study is extremely important because it captures the most recent trends and discoveries in the dynamic field of facial re-identification enabled by deep learning. It is a timely and invaluable resource for researchers, practitioners, policymakers, and industry stakeholders since it provides a thorough review of cutting-edge research and developing approaches. Staying up to date on the most recent technological innovations is critical in order to fully utilize the potential of facial re-identification in surveillance systems. This work not only contributes to a better knowledge of the current state of the art, but it also lays the groundwork for imagining the future trajectory of this transformative technology.

II. METHODS

A systematic and rigorous strategy was used to choose the right selection of papers for the literature review, ensuring the inclusion of relevant and high-quality research. The process is outlined in the following steps, as suggested by [8]:

1. Defining Inclusion and Exclusion Criteria: The publication timeframe, major topic matter (deep learning for facial re-identification in surveillance cameras), and categories of publications (peer-reviewed articles, conference papers, and respectable journals) were all clearly established criteria for paper selection.

2. Comprehensive Database Search: A thorough search was carried out across numerous academic databases, including PubMed, IEEE Xplore, ACM Digital Library, Google Scholar, and Scopus. To
locate relevant research publications, keywords and search phrases included versions of "facial re-
identification," "deep learning," "surveillance cameras," and related topics.

3. Initial Screening: The first search results were filtered based on the title and abstract. Papers that did not match the inclusion requirements were removed from consideration. Duplicate items were also deleted to guarantee that the dataset was unique (Fink, 2019).

4. Full-text Review: The content of each publication was critically examined to determine if it addressed the specific issue of deep learning for facial re-identification in surveillance cameras. Selected papers from the first screening received a full-text examination. Papers that failed to fulfill this criterion were rejected.

The goal of using this thorough technique was to pick a relevant, up-to-date, and high-quality group of articles for the literature evaluation, therefore providing a solid foundation for the study analysis and synthesis. The technique produced the results shown below.

The next part will contain the newest developments and cutting-edge technologies within this subject.

III. RESULTS

In contemporary times, there has been a notable proliferation of deep learning technologies in several computer vision areas [27-31]. The aforementioned domains include face recognition (32, 33), object recognition (34, 35), object detection (36-38), and several others. Technology is a powerful tool that has the ability to do computer vision jobs in a comprehensive manner, eliminating the requirement for manually crafted features [39-53]. The emergence of a significant body of literature has been driven by the accomplishments of this technological progress, leading to the utilization of deep learning
approaches in addressing the issue of person ReID. This section presents a succinct overview and synthesis of diverse tactics.

3.1: Model to Identify the face
Similar to prior recognition methodologies that rely on deep learning and multi-class classification, these deep models for identification regard the person ReID issue as a task of categorizing the attributes of input person image graphs in order to assign corresponding labels. Figure 3.1 depicts the key underlying structure of the identification model. In one of the studies, the researchers propose the utilization of a fusion feature network (FFN) to augment the features of a convolutional neural network (CNN) [10]. This methodology entails the use of a variety of manually designed attributes, including color histogram features and texture features, with the convolutional neural network (CNN) features. In the backpropagation phase, the convolutional neural network (CNN) features are constrained by the wide range of manually crafted features.

The training of the complete network is achieved by the use of softmax loss. To mitigate intra-personal discrepancies and promote inter-personal distinctions, a study [11] proposes the utilization of a hybrid deep architecture for person ReID. The architectural design incorporates low-level descriptors, such as color histograms and Scale-Invariant Feature Transform (SIFT), into Fisher vectors. Following this, the Fisher vectors are integrated with a deep neural network in order to provide nonlinear features that aptly depict pedestrian images. The authors [12] utilize a Convolutional Neural Network (CNN) in their research to provide extensive deep feature descriptors that can be applied to various areas of person Re-identification (ReID) datasets. Figure 3.1 illustrates the REID process as outlined in references [11] and [12].

![Identification Model](image)

**Fig 3.1: Identification Model**

3.2: Model to Verify the recognition
The purpose of the verification model is to evaluate the similarity between two image graphs with the aim of determining if they portray the same pedestrian or not. Figure 3.2 depicts the underlying structural framework of the verification model. Within the realm of person ReID, the verification model is frequently addressed as a binary-class classification issue, as evidenced by prior research [13-15]. The researchers in [13] originally proposed the verification model as a potential resolution to the person ReID problem. The authors propose a novel methodology known as the Filter Pairing Neural Network (FPNN) to effectively tackle various challenges associated with geometric transformations, imagometric...
variations, misalignment, background clutter, and occlusions. This approach integrates max-out pooling layers and patch-matching techniques. The authors in [18] proposed a deep network referred to as "siamese" for the purpose of metric learning. The architectural design encompasses the utilization of three discrete convolutional networks that function on three unique and non-overlapping areas inside the two images. The generation of the deep descriptors for the two input images is achieved by the utilization of fully linked layers. Afterwards, the cosine function is utilized to compute the distance between the two resultant descriptors. In conclusion, the cosine function produces the similarity score. The PersonNet concept was proposed in [15] as a solution for person ReID, building upon the two previously described endeavors. The model utilized the patch matching layer provided by [17] in order to capture the local interaction between patches. Moreover, the use of a deep architecture employing smaller convolution filters confers benefits in terms of augmenting the total depth of the design. One possible drawback of the verification models outlined above is their modest network depths, which restricts their ability to extract profound discriminative features.

Furthermore, the verification model necessitates the generation of image pairings as input and depends exclusively on weakly labeled datasets [23], so compromising the efficacy of the training process. It is important to highlight that the combination of the identification model and verification model has produced promising results in the domain of person ReID. Researchers in [70] proposed an initial approach that incorporates verification and identification losses to train a neural network for the specific task of face recognition. The researchers in [19] proposed a unique technique that combines the verification and identification losses during the training of the Caffe-Net for person ReID, with the aim of maximizing the advantages offered by the identification and verification model. The extant literature offers a comprehensive examination of the advantages and disadvantages linked to both theoretical frameworks. In contrast to the face recognition network, the preference for utilizing cross-entropy loss over contrastive loss is observed. Before this, the process of embedding might potentially be subjected to dropout regularization. Figure 3.2 provides a visual representation of the facial verification model's verification process, as described in references [12] and [13].

Figure 3.2: Facial Re-recognition model verification
3.3: Deep Model Based on Section-based strategy

The methodology outlined in citations [20-37] utilizes a part-based approach in order to get discerning deep characteristics for the purpose of person re-identification (ReID). The researchers utilized a global convolution layer to extract global convolution characteristics in their study [12]. Following that, the researchers divided these characteristics into four equally sized separate branches to get part-based deep features. In the final analysis, the researchers integrate the global and part-based feature vectors to get the ultimate deep features. Similar to the methodology employed by references [47] and [48], the division strategy is utilized. Nevertheless, the network undergoes optimization by the utilization of multi-classification losses. The authors of the research resized the input imagegraphs to dimensions of 128×64 pixels [39]. Following this, the imagegraphs are partitioned into three intersecting sections, with each section having dimensions of 64×64 pixels. Afterwards, the three separate branches are integrated into the three overlapping components, and then a completely linked layer is employed to combine the deep features obtained from these branches. Figure 3.3 depicts the section-based technique.

IV. DISCUSSION

Enhancing face re-identification systems has seen a notable upsurge in research and innovation in recent years. The increasing need for more precise, dependable, and effective techniques in public safety, law enforcement, and surveillance is what's behind this spike. These advancements signify a major step forward in tackling the intricate problems of identity recognition and verification in a variety of contexts. The field of Identification Models, which are the foundation of facial re-identification systems, has seen one significant development. New methods have been developed by researchers to increase person ReID accuracy. For example, a few authors developed the idea of Fusion Feature Networks (FFN), which integrates hand-crafted characteristics like color histograms and texture data with Convolutional Neural
Network (CNN) features. In the end, this fusion technique improves inter-personal differences and decreases intra-personal variations, resulting in more accurate recognition. Hybrid deep architectures have also been developed, which combine deep neural networks with low-level descriptors like color histograms and SIFT to provide non-linear features that accurately characterize pedestrian images. Reliable performance of recognition systems under a variety of settings is made possible by this feature integration.

Deployments of Verification Models, which evaluate image similarity between pairings, have also advanced significantly. Accurately matching and comparing face characteristics depends on these models. The Filter Pairing Neural Network (FPNN), presented by [36], is one such advancement. FPNN is a potential method for robust verification since it successfully handles issues with geometric changes, imagetric fluctuations, misalignment, background clutter, and occlusions. A few researchers introduced Siamese deep networks, which have shown to be successful in precisely determining the degree of similarity between images—a crucial component of facial re-identification. Furthermore, addresses the drawbacks of shallow network depths by combining patch matching layers with a deep architecture that uses smaller convolution filters, enabling the extraction of more discriminative features.

The application of region-Based Strategies, which divide images into portions and extract characteristics from each region, is another noteworthy trend in recent advancements. This method helps to provide a more thorough grasp of a person's look, particularly in situations when there are obscured or incomplete views. To extract global convolution features, for example, a few of the authors used a global convolution layer. These characteristics were then divided into several branches to generate part-based deep features. This fusion of part-based and global information improves the system's individual recognition accuracy. Furthermore, using multi-classification losses to section-based techniques helps to improve person ReID by refining deep characteristics.

The dynamic nature of facial re-identification research is shown by these recent advancements, which emphasize on feature extraction, verification accuracy, and overall system resilience. A viable path forward for the discipline is the combination of section-based techniques and identification and verification models. With its ability to accurately and consistently identify people in a variety of situations, face re-identification is set to become an increasingly important tool in maintaining public safety and security as technology advances.

**CONCLUSION**

In summary, there has been a lot of advancement in the field of face re-identification due to the growing need for reliable and effective systems for recognition and verification in the fields of law enforcement, surveillance, and public safety. Deep learning has revolutionized the way we automatically learn and extract complex facial characteristics and patterns, and it has played a vital role in changing the face recognition environment.

Considerable advancements in facial recognition systems have been made possible by deep learning, which places a strong focus on deep neural networks that are inspired by the structure of the human brain. Convolutional Neural Networks (CNNs) have become an indispensable instrument for the quick and precise interpretation of facial data. This change has improved face recognition's precision and consistency, making it capable of operating well in difficult scenarios with changing illumination, angles, and occlusions.
A branch of computer vision called facial re-identification attempts to solve the urgent problem of person recognition and verification in a variety of camera angles, ambient conditions, and real-world scenarios. The importance of face re-identification has been highlighted by the widespread use of surveillance cameras and the need for improved security and crime prevention. For law enforcement organizations, it is an invaluable instrument that helps with both the prosecution of criminal cases and the improvement of general security protocols. Moreover, its use in congested public areas such as stadiums and airports aids in the prompt identification and reduction of threats, effectively tackling growing security apprehensions.

It is impossible to overestimate how revolutionary deep learning may be in solving the problems associated with facial re-identification. It is revolutionary in its capacity to manage the inherent variety in facial appearances caused by variables like lighting [23], viewing angles [34-38], and expressions [12]. Furthermore, Deep Learning models provide robust generalization and flexibility, guaranteeing reliable performance in intricate and challenging monitoring scenarios.

The most current advances in facial re-identification technology, especially those enabled by Deep Learning, have been clarified by this review. It is a useful and current tool for scholars, professionals, decision-makers, and industry participants. Taking full use of facial re-identification in surveillance systems requires keeping up with the newest technology developments. Moreover, this study offers the groundwork for imagining the future course of this revolutionary technology in addition to improving our comprehension of the state of the art now.

REFERENCES


