ISSN 2393 - 9850
Pages : 47-52

Face Recognition and Image Retrieval: A Comprehensive Analysis for Event-Based Applications

Nihal Muhammed Riyaz
College of Engineering, Chengannur

Nived G
College of Engineering, Chengannur

Noel Paul George
College of Engineering, Chengannur

Shaniya N
College of Engineering, Chengannur

Shabana Mol S
Assistant Professor,
College of Engineering, Chengannur

Abstract

Face recognition and image classification are increasingly prevalent in event photography, enabling the swift identification of guests' photos. This research investigates various techniques for face detection and recognition, particularly in contexts where users upload a portrait image to access all associated photographs. The technology identifies attendees by generating embeddings and comparing them to a collection of event images, eliminating the need for manual searches.To improve the user experience, when users upload their portrait photos, the system is able to take out the unnecessary background if needed. Using Convolutional Neural Networks (CNNs) for face recognition and Generative Adversarial Networks (GANs) for object removal and inpainting, the technology creates smooth and natural-looking backgrounds that enhance the quality of the photos. CNNs are especially good at finding key details in images, making face recognition more accurate and efficient. By combining CNN-based face recognition with object removal and inpainting, the system offers an easy-to-use solution, helping guests quickly find and enjoy their event photos without any distractions.

1 Introduction

The advancements in computer vision and deep learning have generated a great deal of interest in face recognition and image classification. These technologies are essential for automating security and personal identity-related processes. This paper reviews several face recognition and classification techniques with a focus on real-world applications.

Helping attendees of events,like guests at weddings,locate their images is one noteworthy application that was covered. Attendees have the option to upload a portrait image to the system, and event organizers collect a large quantity of photos. After processing the portrait image and comparing it to the event photos, the software extracts all of the pictures that feature the attendees.The same techniques can be utilized in areas of surveillance and security where we could track individuals using their facial images even from huge datasets of images.

Face embeddings are generated by the system after face recognition-based facial extraction from event photographs, which may be compared with the uploaded portrait image.Thus it eliminates the need for attendees to sort through large collections in order to view their photographs but can quickly access them.

The most recent approaches to classifying photos use deep learning techniques to extract feature representations from the images. These techniques include generative adversarial networks (GANs) and convolutional neural networks (CNNs). The methods used in picture improvement, object removal, inpainting, and identification are also covered in this paper.

2 Literature Review

2.1 Face Detection Techniques

Detecting faces is an essential initial stage in recognizing faces. An approach to reduce face detection time in images and videos by focusing on a Search Area Boundary (SAB) instead of scanning the entire frame can be employed to reduce computational complexity and processing time, with an accuracy drop of no

Received : 02 October 2024
Revised : 18 November 2024
Accepted : 30 November 2024

more than 7% [1]. The method is implemented using the DLib library, a popular tool for face detection, and is tested on the DeepFake dataset. The method is effective for videos containing one or more faces and proves efficient in both detection time and accuracy compared to traditional methods. The paper also addresses challenges like dynamic face appearances and updates the SAB when new faces emerge, ensuring continuous performance throughout the video sequence [1].

An improved face detection algorithm based on Retinaface can be employed in view of the insufficient use of the semantic information of high-level feature maps under the existing face detection algorithms [2]. It uses the backbone feature extraction network ResNeSt-50 to extract features, combine the feature pyramid CE-FPN to perform multi-scale fusion of feature maps of different scales.Then information processed by the feature pyramid is transfered to the context module, and finally passing the training and loss function Soft non-maximum suppression to realize face detection and feature point positioning. Experimental results show that the average accuracy of this model on the Wider Face dataset is 94.8%, which is better than the original Retinaface model [2].

Traditional face detection methods often need help with complicated images involving multiple faces, varied expressions, orientations, and lighting conditions. A methodology to outperform this difficulty is introduced using the ResNet , a deep residual learning framework. [3] The paper discusses the implementation details of the ResNet model in Python and how it has been optimized for multi-face detection.With real-world applications in mind, the high accuracy of 99.73% achieved in this method implies a reduced likelihood of false negatives and significantly improved ability to detect multiple faces in real-world scenarios.

2.2 Deep Learning Architectures for Face Recognition

Additionally, Eigenfaces, Fisherfaces, and Local Binary Pattern Histograms (LBPH) are Integrated for face recognition tasks [4]. While these traditional techniques were useful, CNNs outperformed them in terms of accuracy and flexibility. The paper demonstrates how OpenCV and CNN-based models can work together for real-time face recognition systems, showing significant improvements in terms of performance and usability.

2.3 Support Vector Machines (SVM) vs. CNNs

Traditional machine learning methods such as Support Vector Machines (SVMs) have also been used for image classification tasks. While comparing SVM and CNN in classifying images of cats and dogs, CNN outperformed SVM in terms of accuracy, with CNN achieving 89% accuracy while SVM only achieved 61%. CNNs were more capable of handling complex patterns in images due to their deep learning capabilities, whereas SVMs struggled with high-dimensional data [5].

This comparison highlights that while SVMs can still be effective for simpler binary classification tasks, CNNs are better suited for more complex image classification problems, including face recognition.

2.4 Face Image Clustering

The usually followed methods for face image clustering involves the K-means clustering,Principal Component Analysis(PCA),Linear Discriminant Analysis (LDA) etc.But these methods have a short-coming,they need to use the number of clusters as input. When the number of clusters is unknown, we have to estimate the number from the data themselves,which we achieves using LGBPHS model and keeping a threshhold value. [6] There are two important parameters in this method, threshold1 and threshold2. They control the clustering of face images and affecting performance of the algorithm in almost every aspect: precision, speed and the clustering results. The thresholds are obtained in this way: after computing all the Euclidean distances between every two face images of different people in Yale Face database, we fInd that the distances agree with Gaussian distribution, and assign the minimum of the distances to threshold1. The probability that the distance between two face images of different persons is lower than threshold1 is very small, so if the distance of two face images is lower than threshold1, the two can be seen as in the same class.In this method, at first, the face images are clustered into many clusters while each cluster has relatively few images, in the second stage, clusters formed in the first stage are merged in this way: for each cluster A, find its nearest neighbor cluster B. If the distance between A and B is less than a threshold2, A and B will merge, otherwise they will not. Thus, the total number of clusters may decrease. When all clusters finish this operation, go back to the second stage until the total number of clusters stays unchanged [6].

2.5 Transfer Learning in Image Classification

Transfer learning has proven to be highly effective for image classification tasks. Models like ResNet50 and InceptionV3 were used with transfer learning to classify images [7], achieving high validation accuracy (up to 94.26%). The study emphasized how leveraging pre-trained models for specific tasks like face recognition or image classification reduces the time required for training and enhances accuracy when working with smaller datasets.

Transfer learning enables the application of pre-trained models on tasks similar to those they were initially trained on. For example, fine-tuning these models for face recognition allows them to learn domain-specific features, thus improving performance on tasks like face matching.

2.6 Vector Databases and Face Embeddings

Face embeddings represent faces as vectors in a high-dimensional space, facilitating the comparison of faces based on distance metrics such as cosine similarity or Euclidean distance. These databases allow for fast querying of similar faces, making them ideal for large-scale face recognition systems [8].

Vector databases such as Milvus, Weaviate, and Qdrant are optimized for handling high-dimensional data, allowing for quick and accurate face retrieval [8]. These databases work by converting unstructured image data into vectors, which can then be compared and matched, providing the foundation for face recognition systems that require high performance and scalability.

2.7 High Resolution Face Image Enhancement via GAN

The problem of low-resolution facial images can be addressed using deep learning a GAN-based solution [9], focusing on enhancing the quality of facial images for biometrics, surveillance, and recognition of faces. The study focuses on a modified GAN-based model called HRFGAN (High-Resolution Face Image Enhancement via Generative Adversarial Networks). By tailoring the model to the domain of face images, they achieved superior performance compared to general-purpose models.

The Generative Adversarial Network consists of a generator network and a discriminator. To align with the model's input and output specifications, the face images from an FFHQ dataset are downscaled using bicubic interpolation, thus creating a new dataset with pairs of upscaled and downscaled images. The downscaled images serve as the input to the generator during the training process and aim to produce an upscaled version which is evaluated by the discriminator [9].

The generator network contains convolutional layers, activation functions such as Parametric Rectified Linear Unit (PReLU) and Leaky rectified linear unit (LReLU), a series of 12 residual networks and pixel shufflers, which performs the task of upscaling the image.
The discriminator network consists of a series of convolutional layers, Leaky Rectified Linear Unit (LReLU) activation functions, and fully connected layers. The discriminator's job is to evaluate the upscaled images from the generator, comparing them to the actual high-resolution images from the dataset, using various evaluation functions such as Peak Signal-to-Noise Ratio (PSNR). The visually pleasing results of HRFGAN demonstrate its ability to enhance facial details and improve the overall quality of low-resolution face inputs.

2.8 Object Removal and Inpainting from Image using Combined GANs

A breakthrough in image manipulation has been made through GANs, particularly with tasks related to deleting unwanted objects in images or filling in the missing parts. Previously, patch-based or exemplarbased techniques were dominant [9]. Before the emergence of GANs, these previous methods had been extensively worked on manually and did not perform well for larger areas and complex backgrounds. The entity drove the ability to automatically perform object removal and obtain realistic backgrounds that blend smoothly. This was an opportunity to replace the drawbacks of available methods [10].

One important approach uses two GANs to remove objects and fill in the background. The first GAN removes the object from the image, while the second GAN generates a background that fills the empty space. This process happens in one step, without needing separate object detection methods. Tested on the CityScapes dataset, this method has shown good results in removing objects like cars and people from street scenes, making the images look natural [10].

Many datasets are common in training and testing these models. One of them is the CityScapes Dataset which is based on urban scenes; it can also be used to test object removal within a city setting. A dataset

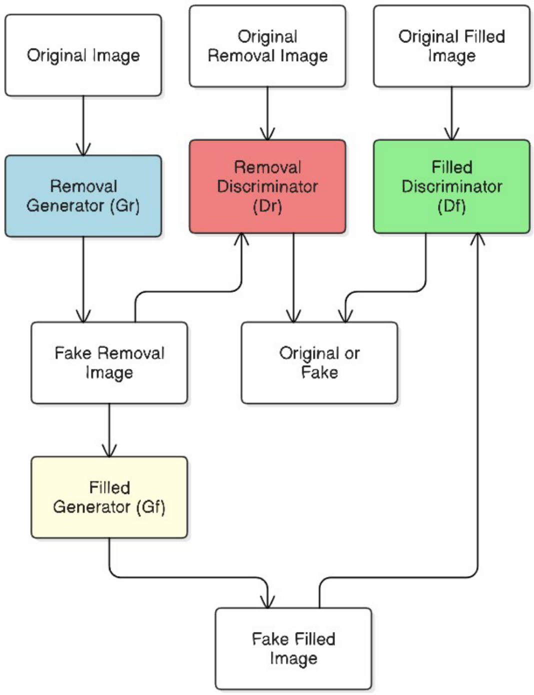

Figure 1: Object removal and background filling using two GANs

with many everyday objects is COCO [11] and it is among the most used datasets for general object removal. Places2 is another dataset; it contains millions of images from different settings which is beneficial for training models to inpaint many different types of scenes. Testing these networks on such qualifying datasets will guarantee the desired performance of GANs in varied contexts [11].

It, however still comes with some challenges in removing larger objects or handling scenes. Most times, small objects are removed with no traces, but when the object in question is larger or the background is complicated, the results may sometimes look less realistic. Further work might apply to improve how well the background is recreated or let the user guide the object removal process for better results. With further development, GANs could be even more useful for applications such as image editing, video processing, and autonomous driving.

Table 1: Comparison of face recognition techniques on various performance metrics.

Technique	Accuracy	Computational Complexity	Dataset Used
Haar Cascades	80%	Low	LFW
HOG + SVM	85%	Moderate	FDDB
CNN-based Models	95%	High	CelebA, VGGFace
GAN-based Techniques	90-93%	High	Custom datasets

3 Significance

These face recognition and classification techniques becomes more versatile when it comes to a dynamic environments where computer vision is inevitable to track the activities.It doesn't confine to attendee or guest satisfaction but can play huge roles in serving administrative and judiciary agencies in citizen tracking, criminal identification and in many other fields of security.

In situations that require real-time results with minimal latency,these face recognition systems can be integrated with edge computing techniques.Thus, any delays associated with transferring the data to centralized servers can be eliminated at places of conferences,concerts etc.The dependency over stable internet connectivity for uninterrupted functionality as in the case of cloud based systems can be reduced by edge computing since most of the processing is handled locally.It also minimizes the risk of data or privacy breach as well as loss of image or video quality as only minimal data are sent through the network.

4 Challenges

There could still arise challenges in AI-based face recognition and classification systems as there can be pose variations in the image to be analyzed when compared against the reference images, such as the faces appearing under different lighting conditions, wearing mask also might interrupt the accuracy of image detection. Also, changes in facial expressions can modify the facial features leading to lower recognition performance.

Images with very low resolution can significantly impact accuracy as they cannot be upscaled to a considerable level. Improving the preprocessing models like noise reduction, adopting data augmentation strategies, etc. might help resolve this issue somewhat.

While working with face detection models, persons in the background might be detected accidentally along with the foreground ones, which could lead to confusion and inaccurate results.

5 Conclusion

To conclude, the precision of face recognition systems has considerably improved by incorporating methods like transfer learning and face embeddings. One key achievement in this area is the implementation of vector databases, which make it easier to store high-dimensional face embeddings quickly and efficiently on a large scale. By arranging these embeddings as vectors, vector databases offers quick searches, along with facial recognition even with large datasets used in security or cloud services at real-time.Also, a successful method is implemented for face clustering tasks without using cluster number as a prior where the algorithm automatically produces cluster numbers and the clusters.

Moreover, generative models have improved the quality of images and made tasks such as object removal and image inpainting more manageable. Despite the existing challenges, especially with managing larger or more intricate data, ongoing advancements in algorithms and database technologies are broadening the capabilities of face recognition and image classification systems. These advancements show potential for the future, creating opportunities for additional applications and advancements in the field.

References

[1] Qadir, Jihad Anwar, Shayan Ihsan Jalal, Samer Kais Jameel, and Abdulbasit Al-Talabani. ”Improved Time-Saving Face Detection in Video.” In 2023 9th International Engineering Conference on Sustainable Technology and Development (IEC), pp. 1-6. IEEE, 2023.

[2] Chen, YanFei, ZeCheng Wu, HongSheng Peng, Chao Zhou, Zhangchen Yan, and Yuliang Huang. ”Face detection algorithm based on improved Retinaface.” In 2021 China Automation Congress (CAC), pp. 4542-4547. IEEE, 2021.

[3] Kumar, Munish, Monali Gulhane, Sandeep Kumar, Harshita Sharma, Rashmi Verma, and Deepika Verma. ”Improved Multi-Face Detection with ResNet for Real-World Applications.” In 2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART), pp. 43-49. IEEE, 2023.

[4] Pandey, Nandita, Poonam Kumari Yadav, and K. P. Arjun. ”Face recognition system using computational algorithms.” In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 1739-1744. IEEE, 2022.

[5] Kausar, Afrah. ”Comparitive Analysis of Support Vector Machine and Convolutional Neural Network in Image Classification.” In 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), pp. 1-6. IEEE, 2023

[6] Deng, Qiang, Yupin Luo, and Junfeng Ge. ”Dual threshold based unsupervised face image clustering.” In 2010 The 2nd International Conference on Industrial Mechatronics and Automation, vol. 1, pp. 436439. IEEE, 2010.

[7] Rajendran, Adithya, S. Abhishek, Adarsh Krishnan, Nived Krishnan Ramesh, and T. Anjali. ”CogniVision: Deep Learning Approaches for Image Recognition and Classification.” In 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), pp. 511-516. IEEE, 2023.

[8] Kukreja, Sanjay, Tarun Kumar, Vishal Bharate, Amit Purohit, Abhijit Dasgupta, and Debashis Guha. ”Vector Databases and Vector Embeddings-Review.” In 2023 International Workshop on Artificial Intelligence and Image Processing (IWAIIP), pp. 231-236. IEEE, 2023.

[9] Venkidesh, K., D. Akhilaraj, and Joseph Zacharias. ”High-Resolution Face Image Enhancement via Generative Adversarial Networks.” In 2023 Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS), pp. 1-6. IEEE, 2023.

[10] Pyo, Jeongwon, Yuri Goncalves Rocha, Arpan Ghosh, Kwanghee Lee, Gungyo In, and Taeyoung Kuc. ”Object removal and inpainting from image using combined GANs.” In 2020 20th International Conference on Control, Automation and Systems (ICCAS), pp. 1116-1119. IEEE, 2020.

[11] Wang, Haoyu, and Lulu Guo. ”Research on face recognition based on deep learning.” In 2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture (AIAM), pp. 540-546. IEEE, 2021.