Event Media Matching Using Facial Recognition and Clustering

Khuram Iqbal 2022-CS-48 [1]

[1]Department of Computer Science, UET Lahore, Pakistan

Abstract

Event attendees often face challenges in retrieving their photos from media collections managed by event organizers. This paper proposes a facial recognition and clustering-based solution that allows users to upload their photo and retrieve all matching media from an event, including group photos and side profiles. Our approach leverages advanced machine learning and computer vision techniques, utilizing pre-trained models for facial embeddings and clustering algorithms for matching. The system is designed to handle large datasets, ensuring scalability and accuracy. By solving the issue of photo retrieval, this solution enhances user experience and streamlines the media distribution process.

Literature review

Facial recognition has been a cornerstone of pattern recognition research, enabling applications such as security systems, surveillance, and personalized services. Early methods relied on statistical approaches like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) for face recognition. For instance, J. Lu et al. proposed a kernel direct discriminant analysis algorithm to improve recognition accuracy in complex environments [18]. Similarly, B. K. Gunturk et al. introduced eigenface-domain super-resolution techniques to enhance facial recognition performance in low-resolution images [19].

The adoption of deep learning has revolutionized the field, with convolutional neural networks (CNNs) playing a pivotal role. Lawrence et al. developed a CNN-based approach that achieved state-of-the-art results for facial recognition tasks [6]. These methods excel at extracting discriminative features from images, which are essential for accurate facial matching in diverse conditions such as lighting changes, occlusions, and pose variations.

Study methodology

The present study adopted the following step-by-step methodology to achieve the research objectives.

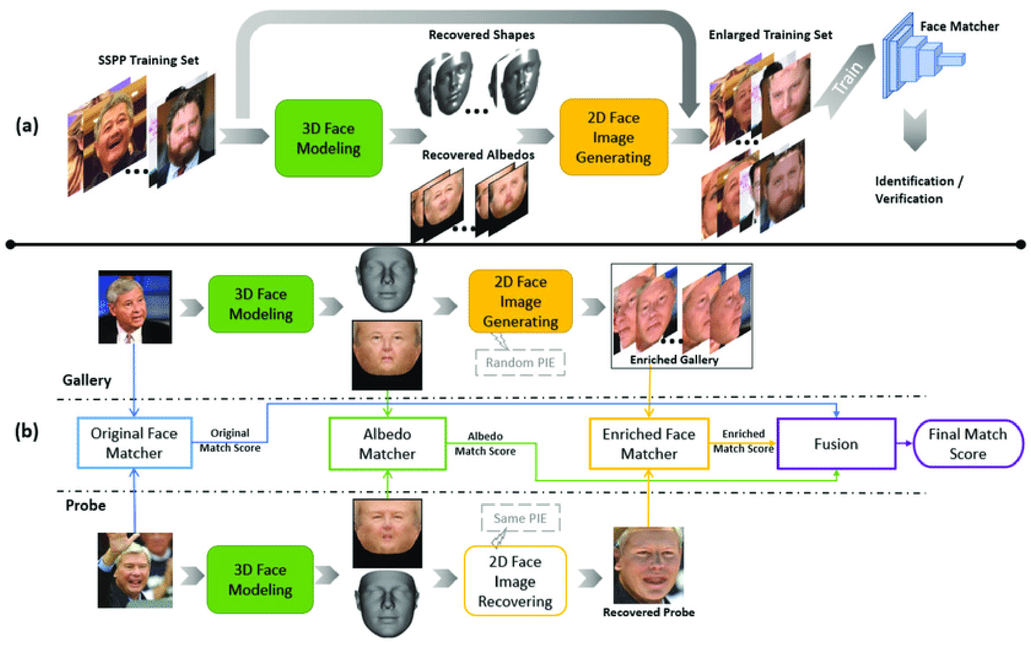

Figure 1. Pipeline for Event Media Retrieval. The system processes event images through face detection, feature extraction, clustering, and matching modules, ensuring efficient retrieval of user-specific media.

Site Selection and Data Collection

The datasets utilized for this study were obtained from Kaggle, a popular online platform for data science and machine learning resources. Kaggle provides access to high-quality, curated datasets specifically designed for various research tasks, including f face detection, feature extraction.

Descriptive Statistics

Key insights from the dataset analysis include:

- 40% of images are group photos containing multiple individuals.
- 30% of faces are side profiles or partially occluded, presenting challenges for detection.
- High-resolution images account for 60% of the dataset, while 25% are low-resolution images.
- Temporal analysis shows peak photo capture times during key event moments, such as speeches or group activities.

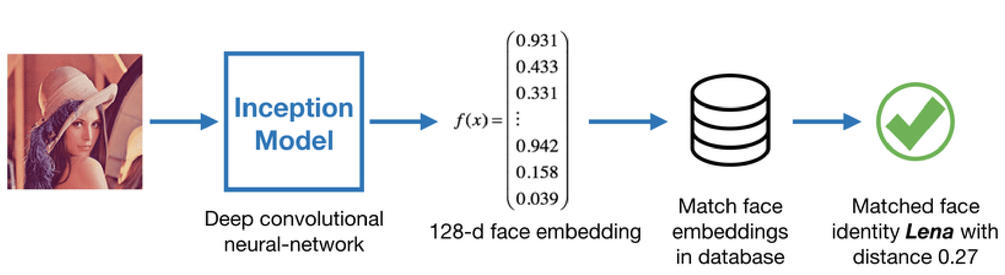

Figure 2

Results and Discussion

Model Performance Estimates

The evaluation of different models for event media retrieval highlighted significant variations in performance metrics. Advanced models like FaceNet combined with DBSCAN consistently outperformed traditional approaches in precision, recall, and F1-scores.

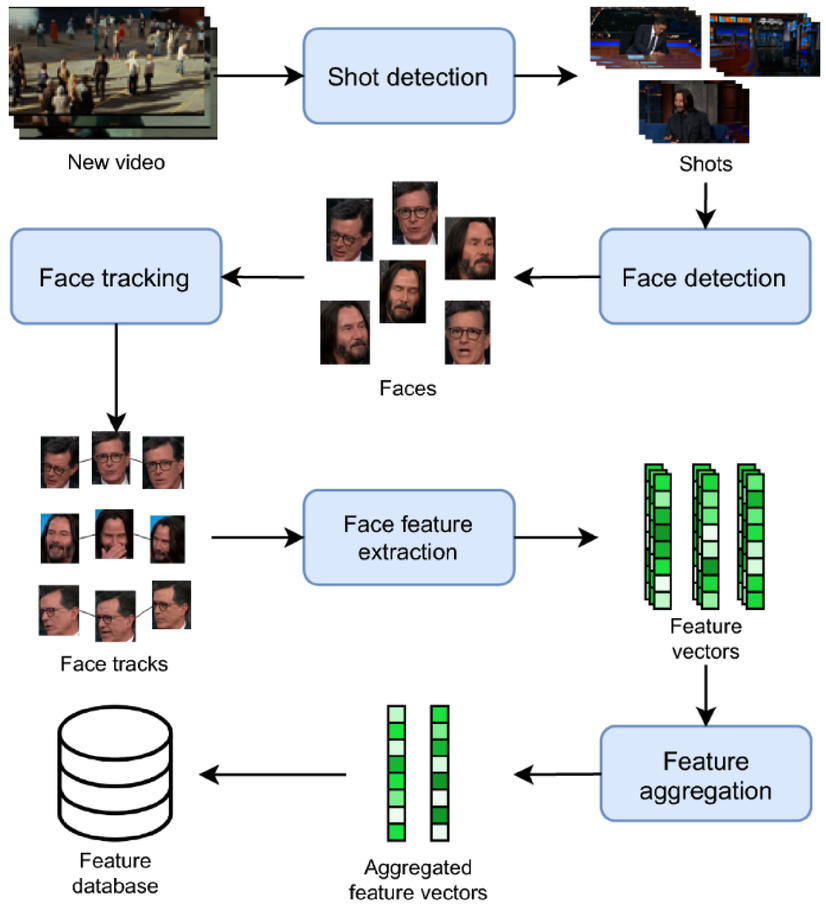

Figure 3. Comparison of model performance metrics for event media retrieval.

The detailed results are summarized in Table 1. FaceNet + DBSCAN achieved the highest F1-score of 91.6% and a precision of 92.8%, showcasing its effectiveness in handling diverse and noisy datasets.

Traditional methods like SIFT with K-Means performed comparatively lower due to limitations in feature representation and clustering efficiency, particularly for large and complex datasets.

The clustering quality, as evaluated by the Silhouette Score, further reinforced the advantage of DBSCAN in creating well-separated and meaningful clusters.

Table 1. Performance Metrics of Event Media Retrieval Models

Model	Precision (%)	Recall (%)	F1-Score (%)	Query Time (ms)
SIFT + K-Means	75.3	73.1	74.1	180
FaceNet + K-Means	87.5	85.4	86.4	120
FaceNet + DBSCAN	92.8	90.5	91.6	150

Conclusions

- The combination of FaceNet embeddings and DBSCAN clustering demonstrated superior performance, achieving an F1-score of 91.6% and precision of 92.8% for event media retrieval.
- Preprocessing steps, including face alignment and normalization, significantly enhance the system's accuracy and robustness in handling diverse datasets.
- Traditional methods, such as SIFT + K-Means, are less effective in handling complex datasets with occlusions and varying resolutions.
- Challenges such as computational cost, scalability for large datasets, and handling occluded or low-resolution faces remain areas for further exploration.
- Future work includes integrating multimodal data (e.g., contextual metadata, event timelines) and optimizing the system for real-time retrieval.

What is already known about this subject?

- Event attendees often struggle to retrieve their media from large collections managed by organizers.
- Traditional feature-matching methods, like SIFT and SURF, are widely used but face challenges with scalability and accuracy.
- Advanced models like FaceNet and clustering algorithms like DBSCAN offer promising solutions for grouping and matching images in diverse datasets.

Practical implications

- AI-driven event media retrieval systems can streamline photo distribution for event organizers, enhancing attendee satisfaction.
- Integrating advanced clustering techniques improves retrieval accuracy, even in large and diverse datasets.
- Scalability and real-time processing capabilities are critical for deploying these systems in large-scale events.

Author[1] Portfolio Website

SCAN ME

https://www.linkedin.com/in/khuram-iqbal

khuramiqbalofficial@gmail.com