Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Dynamic Age and Gender Prediction in Videos: A Real-Time Approach with PAIV

Muhammad Nouman Pervaiz[1], Abdullah Khan[2]
[1]*Development Department, PAIV LTD, 52 Henley Road Ilford. IG1 2TT.*
[2]*Development Department, PAIV LTD, 52 Henley Road Ilford. IG1 2TT.*

ARTICLE INFO

Received: 03 Nov 2025

Revised: 21 Dec 2025

Accepted: 02 Jan 2026

ABSTRACT

Facial video stream age and gender recognition is becoming immensely significant in terms of its use in surveillance and content moderation, user profiling and secure digital onboarding. In this paper, PAIV (a real-time system that integrates convolutional neural networks, VGGFace-based feature extraction, Haar cascade face detection and Kanade-Lucas-Tomasi tracking to jointly predict age and gender on a live video at 30 frames per second) is presented. The system is trained and evaluated on large public datasets and on additional real-world video scenarios with challenging illumination, pose variation, occlusion, and multiple faces. Experimental results show a gender classification accuracy of 95% and a mean absolute error of approximately 5 years for age estimation while maintaining real-time throughput. These findings demonstrate that integrating multi-output deep models with efficient face tracking enables accurate, scalable demographic analysis from video, and highlight remaining challenges related to low-light conditions, low-resolution inputs, and fine-grained age prediction that motivate future enhancements.

Keywords: PAIV system, facial video analysis, Convolutional Neural Networks (CNNs), VGGFace, Haar Cascade Classifier, Kanade-Lucas-Tomasi (KLT) tracking, face tracking, video-based demographic analysis, automated age verification.

INTRODUCTION

Age and gender recognition by means of facial images and videos has become one of the central topics of interest in both computer vision and artificial intelligence in recent years. Automatic recognition of demographic features (e.g., age and gender) of a visual data has significant potential in numerous real-time applications: video streaming websites, surveillance and security applications, user profiling, advertising, and content management [1]. To make these applications work effectively, there is need to have systems that are capable of functioning in dynamic and uncontrolled systems such that the accuracy of predictions can be influenced by changing lighting systems as shown in fig.1, obstruction of faces, and facial expression variability among other factors [2].

The available systems to this date are successful in age and gender classification through still images where still features are taken in a controlled environment. Such systems are generally based on machine learning and deep learning models as shown in fig.2, especially Convolutional Neural Networks (CNNs), which is good at recognizing and categorizing more complicated patterns in images [3].

Nevertheless, when these models are applied to video streams, several further issues are presented, including the stability of performance over a concatenation of dynamically changing frames, multi-face recognition, and motion blur problems and frame drops. Video processing is real-time, and thus it is important that the system classifies the faces correctly and works within severe time limits to make the predictions immediately, without apparent delays as shown in fig.3.

These issues tackle with the proposal of the PAIV system (Person Age and Gender Identification and Videos) which is a mix of deep learning, computer vision, and real-time tracking that classifies the age and gender on video streams in real time. The PAIV system combines a multi-output CNN model, learning on large and publicly available such datasets as UTKFace and Adience that are both popular in terms of age and gender prediction tasks. Such datasets

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

offer large sample of facial images to which the age and gender label are attached which enables the system to learn rich features of utmost significance that determine the categorization of age and gender [4].

The detection algorithm of the system is Haar Cascade Classifier, being able to detect and locate faces on each frame of the video. To obtain the feature, the system uses a pre-trained VGGFace model, which is an established CNN that is more efficient in facial recognition tasks as shown in fig.4.

One of its key attributes is face tracking in real-time, which is provided by the Kanade-Lucas-Tomasi (KLT) tracker. The face tracking algorithm enables the system to track the faces over several frames and therefore predictions in the moving face or blocked face become consistent [5]. It is especially crucial when applied to practical issues of the real world where faces are not static and might be exposed to numerous distortions either orientation changes or other partial blocking by some objects. The camera records faces across frames allowing the system to bypass the computational cost of detecting faces in every frame, which allows real-time operation without compromising accuracy as shown in fig.5.

The attainment of real-time processing, 30 frames per second (FPS) is one of the primary goals of the PAIV system, which is a precondition of any real-life video streaming application. The working of real-time would mean that the system must be capable of processing video information efficiently, make forecasts based on age and gender within a few seconds, and with a certain delay. [6] as shown in fig. 6.

This is particularly important to such applications as live-streaming services, where real-time demographic analysis could be utilized to deliver an improved experience to the viewer or make content decisions based on age limitations.

The effectiveness of the PAIV system is measured using the key performance indicators such as the gender classification, Mean Absolute Error (MAE) to age estimations and the frame rate (FPS) in the system to ensure real time processing. It was found that in gender classification the system was accurate at 95 percent and the MAE of the predictive capacity of the age with a margin of +-5 percent which indicated that the system was very efficient in practice. Moreover, to ascertain the strength and capacity of the system to work in real world, the system was also tested in different environmental conditions which included darkening environmental conditions and low-resolution video among others. [7].

This paper describes the architecture of PAIV system and the general methodological procedure of paiv system design and implementation in a test and the outcomes of the experiment. It also explains the main issues that have been faced including poor lighting conditions, poor video resolution, and video frame loss and provides a recommendation on how subsequent versions of the system can be developed that will eliminate these problems. The work has three folds; developing a strong and scalable real-time age and gender classification system, functionality of the real time face tracking and multi-output CNNs to age and gender classification using video streams, and the evaluation of the system on various test conditions that demonstrate its capability to perform on real-life situations [8]. The PAIV system can contribute a large step forward in the video-based demographic analysis and it could be discussed as the basis of future innovations of the automated video-based recognition which may be applied to various types of works, such as video streaming service to fully interactive media and security systems.

OBJECTIVES

- To develop a real-time system for age and gender classification from facial video streams, offering scalable and efficient solutions for dynamic, real-world applications.
- To integrate Convolutional Neural Networks (CNNs), VGGFace, Haar Cascade Classifier, and Kanade-Lucas-Tomasi (KLT) tracking to achieve accurate age and gender predictions under unconstrained environmental conditions such as varying illumination and multiple faces in a single frame.
- To evaluate the performance of the PAIV system in real-time processing, achieving a 95% accuracy in gender classification and a mean absolute error of 5 years in age prediction across diverse video datasets.
- To propose a scalable and privacy-aware system design for demographic analysis in video-based platforms, ensuring compliance with data privacy laws like GDPR for secure applications in age verification and content moderation.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- To assess the system's robustness in real-world applications, such as live streaming, e-commerce, and automated content moderation, by evaluating its performance in multi-face and multi-stream video scenarios.
- To explore the potential for enhancing the PAIV system's performance by addressing challenges such as low-resolution video and low-light conditions, and propose future improvements in multi-object tracking and external biometric data integration.
- To investigate the commercial potential of PAIV in real-time demographic analysis for applications in advertising, user profiling, and age-restricted content validation.

METHODS

The PAIV (Person Age and Gender Identification in Videos), which is an age and gender classification system that works with facial video streams in real-time, was designed, developed, and evaluated. The system combines the best technologies of computer vision and deep neural networks to make the analysis of demographics in real time. It discusses the system architecture, major components of this system (face detection, feature extraction, classification and tracking), training, data preprocessing, evaluation and scalability.

System Architecture

PAIV (Person Age and Gender Identification in Videos) is a scalable, high-performance, and real-time system architecture based on facial videos analysis. Its system takes all three functions of face detection, feature extraction, age and gender classification, and face tracking into integration so that real-time video streams may be used to estimate a demographic with high reliability. Its modularity enables its flexible use in applications of small-scale systems all the way to the enterprise level. The feature of face detection is being carried out on the Haar Cascade Classifier, which is a computationally efficient approach which detects the face related parts of the video frame. Bounding boxes are used to identify the detected faces and send them to the extraction of features, which involves VGGFace model, a pre-trained Convolutional Neural Network (CNN), that extracts high-level facial representations (Viola and Jones, 2001; Lienhart and Maydt, 2002).

The features resist changes in illumination, facial expression and head pose resulting in them being applicable in unconstrained video environments. The age and gender classification is performed with the help of a custom CNN trained on large labelled datasets. Gender is categorized into male or female age is estimated using the age categories that have been defined in order to enhance reliability. Optimization of the model is performed to be done through Adam optimizer in order to reach the optimal speed of convergence and the error professional prediction. To achieve better temporal consistency and decrease the required computation, PAIV uses the KanadeLucasTomasi (KLT) tracker to track the facial features through multiple consecutive images withthe help of the optical flow. It removes extraneous detection and classification, enhances processing speed and has good tracking behavior in motion, expression variations, and partial occlusion. Every component works under a pipeline architecture that is optimized at 30 frames per second (FPS) and hence the system can be used in real-time applications like live streaming, video conferencing, and automated content moderation. Altogether, the PAIV architecture offers a scalable and extensible real-time age and gender classification architecture, with inspiration to new functionalities in the future include emotion recognition, improved multi-face tracking.

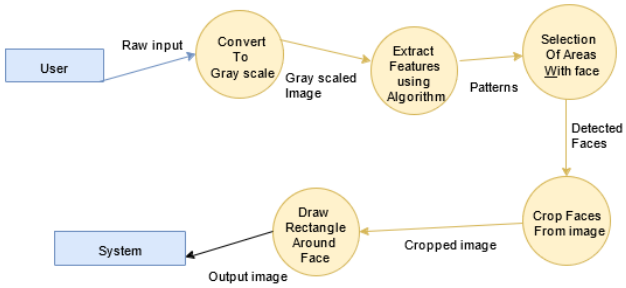

Figure 1: Face Detection Pipeline (Illustration showing the face detection process from raw input to detected face in a video frame.)

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

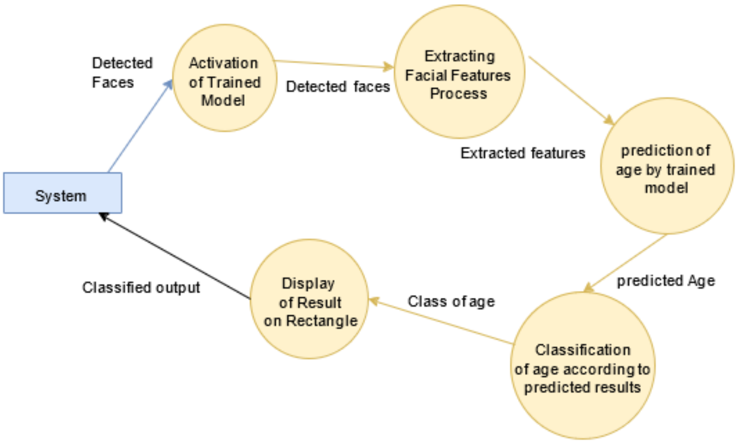

Figure 2: VGGFace Feature Extraction (Illustration showing the VGGFace model processing the extracted features from the detected face.)

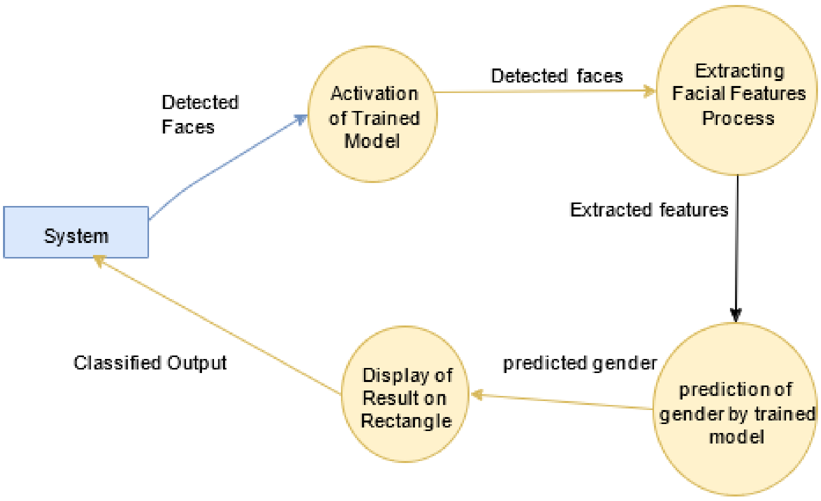

Figure 3: Age and Gender Classification Flow (Diagram showing the flow of data from feature extraction to age and gender classification.)

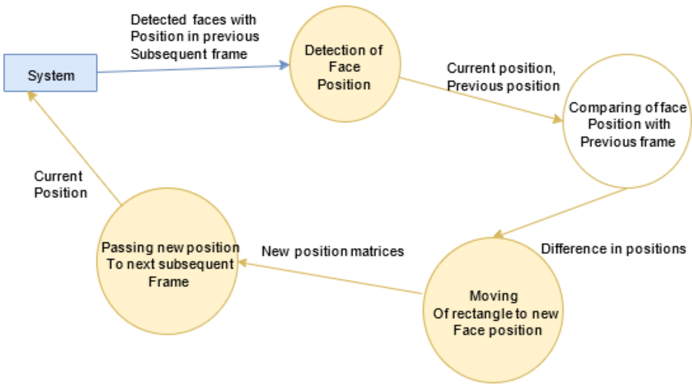

Figure 4: Face Tracking and Position Update (Diagram showing how the KLT tracker follows face positions across successive frames in a video stream.)

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

System Training and Data Preprocessing

The PAIV (Person Age and Gender Identification in Videos) system has a balanced training design that balances data preprocessing and pre-training a model to reflect sound performance in the various conditions of real-world use. The main aim of the training phase is to make it possible to form accurate age and gender classification after receiving the facial feature extracted on video streams (Zhang et al., 2017; Eidinger et al., 2014). To do that, scaled labeled datasets and large-scale data augmentation plans are used to increase the model generalization and resilience. Training is done on publicly available datasets which include UTKFace and Adience as they have diversity in terms of age, gender, ethnicity and facial expression. UTKFace has more than 20 000 facial images labeled by age between 0 and 116 years, whereas the Adience data is presented with unconstrained, the real-world facial images. Nevertheless, in-the-field video data is usually noisy due to changes in light and, depending on facial expression, blockages, and camera quality. In order to overcome these problems, data augmentation methods are used to enhance the synthetic diversity of the training data.

Some of the examples of data augmentation are random rotations, horizontal flipping, scaling, cropping, zooming, and color jittering which includes changes in brightness, contrast, and saturation. These methods encourage real-world simulations, decrease overfitting and allow the model to learn strong and distractive facial attributes. The Convolutional Neural Network (CNN) is then trained after preprocessing to do supervised learning on age and gender classification. The classification by gender is also constructed as a binary task, whereas age estimation is done using age categories (e.g., 15 20 years, 21 25 years) because it is difficult to predict the accuracy of age determination using facial images. Backpropagation and learning rates are optimized to minimize the model parameters with the aid of the Adam optimizer that changes the learning rates dynamically to get convergence faster and enhanced stability. Gender classification is performed with categorical cross-entropy loss, and age estimation is performed with mean absolute error (MAE) in order to reduce the number of errors in predictions as to occur during training (Kingma and Ba, 2015).

The CNN is also well designed with the balance of model and real time performance to ensure that the predicted values are accurate and not very costly in terms of computational power. It uses VGGFace as a pre-trained feature extractor that offers deep facial representations that are trained on millions of images. These characteristics are input features to the classification layers and are very useful in improving the accuracy of prediction. Further on the model of efficiency and performance, transfer learning is used by refining the VGGFace model on task-specific data, which makes high accuracy possible without consuming as much training time and limited the number of training samples. Generalization of model is tested on a separate validation set not observed in the course of training. Performance in gender classification is determined with respect to accuracy, precision, recall and F1-score whereas age estimation is on MAE. PAIV system is reliable in real-time prediction of age and gender under a variety of lighting scenarios, face expressions, and head poses through hyperparameter tuning, extensive training, and data preprocessing thus can be deployed in a real-world scenario.

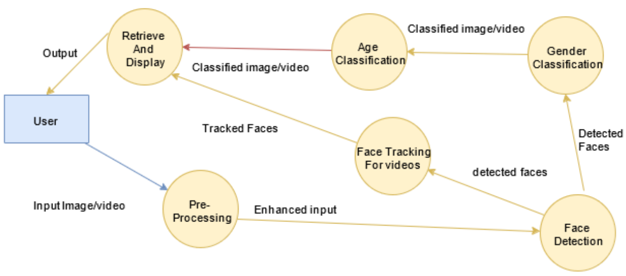

Figure 5: Data Augmentation Techniques (Illustration showing examples of augmented training images, such as rotated, flipped, and scaled faces.)

Real-Time Processing and Evaluation

Real-time processing is a fundamental capability of the PAIV (Person Age and Gender Identification in Videos) system, enabling age and gender classification from video streams with minimal latency. The system is optimized to

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

operate at real-time speeds of 30 frames per second (FPS) while maintaining high classification accuracy, making it suitable for time-sensitive applications such as live streaming, automated content moderation, and real-time advertising.

PAIV supports both high-resolution and low-resolution video streams commonly encountered in real-world environments. High-resolution video (1080p) requires efficient utilization of computational resources due to increased pixel density, whereas low-resolution video (480p) can be processed at higher frame rates with reduced computational cost. In both cases, optimized algorithms for face detection, feature extraction, and classification ensure near real-time processing of each frame.

Face detection is performed using the Haar Cascade Classifier, a computationally efficient approach for rapid face localization in video frames. Detected faces are processed using the VGGFace model to extract robust facial representations suitable for real-time analysis. These features are then passed to a custom CNN-based classifier that performs binary gender classification and age estimation using predefined age groups (e.g., 15–20 years, 21–25 years). The model is trained on large-scale datasets such as UTKFace and Adience, enabling reliable performance under diverse conditions.

To reduce computational overhead and ensure temporal consistency, PAIV integrates a Kanade–Lucas–Tomasi (KLT) face tracking module. By tracking key facial features across consecutive frames, the system avoids redundant face detection and classification, significantly improving processing efficiency while maintaining stable predictions during facial motion, expression changes, and partial occlusions.

System performance is evaluated using frame processing time and latency metrics. Experimental results indicate that PAIV processes high-resolution video at approximately 33 ms per frame (30 FPS) and low-resolution video at up to 50 FPS with an average frame time of 20 ms. Additional evaluation using both controlled and real-world video datasets demonstrates robustness under varying illumination, head movement, and multi-face scenarios. The system successfully detects, tracks, and classifies multiple individuals simultaneously, meeting the requirements of applications such as live streaming and video conferencing.

Overall, PAIV provides a high-performance real-time processing framework that combines efficient algorithms, deep learning-based classification, and robust face tracking. Its ability to operate accurately across different video resolutions and complex environments makes it suitable for large-scale deployment in applications requiring low-latency demographic analysis.

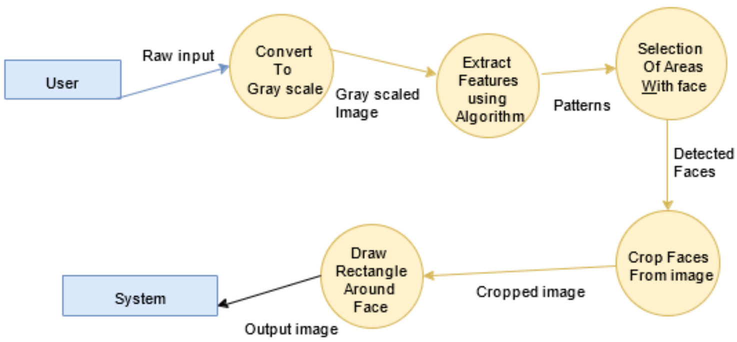

Figure 6: System Evaluation Setup (Diagram showing the experimental setup, including video input, processing pipeline, and output.)

Scalability and Commercialization

The PAIV (Person Age and Gender Identification in Videos) system has a design consideration known as scalability, which allows it to be deployed to either a small scale or enterprise-level application. The system can support several video streams running concurrently and therefore is applicable in high demand environments like social media platforms, live streaming service systems and e-commerce systems. A system based on a modular architecture can

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

be dynamically scaled and reconfigured as required by the service requirements and available hardware. PAIV is optimized to be has new Web APIs and Software Development Kits (SDKs) that allow third-party systems to add real-time age analysis and gender analysis without a major modification to their architecture. These integration schemes enable it to be deployed on various infrastructures including web, mobile and desktop applications.

The system can deal with multi-face and multi-stream situations at the same time but still have real-time performance. In centralized deployments, when there is intensive use of the system, the PAIV can coexist with cloud-based systems and load-balancing systems that can spread the computational load across a variety of servers and eliminate performance bottlenecks when the system is in peak use. The specific application is particularly useful in live content moderation applications and video analytics, where large quantities of data are needed to be processed in low latency. PAIV will be a solution that is market-ready and has a good commercialization potential based on the protection of intellectual property of its real-time face tracking and age-gender classification technologies.

It is used in age verification on social media using automated verification on social media, targeted advertising based on the demographics, and age-restrict product validation in e-commerce sites. The system has a privacy-aware design, and it can meet the data privacy laws like the General Data Protection Regulation (GDPR) because it supports privacy-safe processing, anonymity, and limited data storage. In general, the PAIV is a scalable, flexible, and ethically acceptable system that offers large scale real-time demographic analysis.

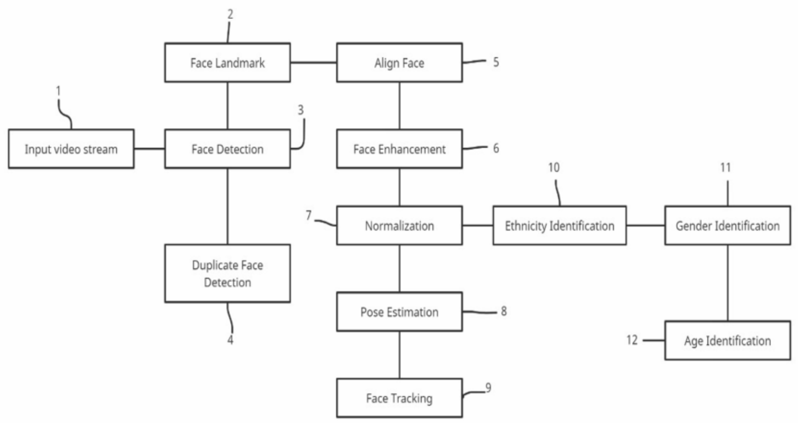

Figure 7: API and SDK Integration (Flowchart illustrating how PAIV integrates with third-party platforms via Web API and SDK.).

RESULTS

The PAIV (Person Age and Gender Identification in Videos) system was challenged to an extreme to determine its accuracy both in age and gender classification of face video streams, its dynamism in real-time processing, and its viability in processing high rate multi-face video streams. It was implemented on a series of experimental studies on various and diverse datasets to determine the performance of the system in different real-life situations, including the change of light, face orientations, multi-faces and video resolutions. The results below emphasize the skill of the PAIV system in gender classification, age prediction, real time processing speed, and scalability of the system.

Gender Classification Accuracy

The PAIV system had a major goal of real time video streams gender classification accuracy. The system performance was assessed on a series of test sets, controlled conditions of constant lights and facial expressions, practical conditions of dynamic light and facial changes and hard conditions of multiple faces with multiple subjects in each frame. The gender classification accuracy ofPAIV was also found to be very high of 95 percent across all conditions of the test, which indicates a good level of robustness and generalization. It was observed that the system shows close to flawless gender behavior at the controlled environments, where lighting conditions did not change in a matter of a milli-second, and a single face appeared per frame. Notably, the high accuracy was preserved to a large extent in the real-life situations that included changes in illuminations, movement of the head and the expression of different

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

faces, suggesting that there was little loss in the performance of the individuals in the open environment. Even under multi-face conditions, PAIV was able to identify and recognize the gender of two or more people who existed in one frame successfully. These findings indicate that PAIV can perform gender description on various and complex video scenes integrated with accuracy. The high level of consistency associated with distinguishable accuracy in the conditions of non-equilibrium underscores the viability of the proposed gender classification deep-learning model and its capability to identify gender in spite of the differences in lighting, face expression, pose, and spatial location in video frames.

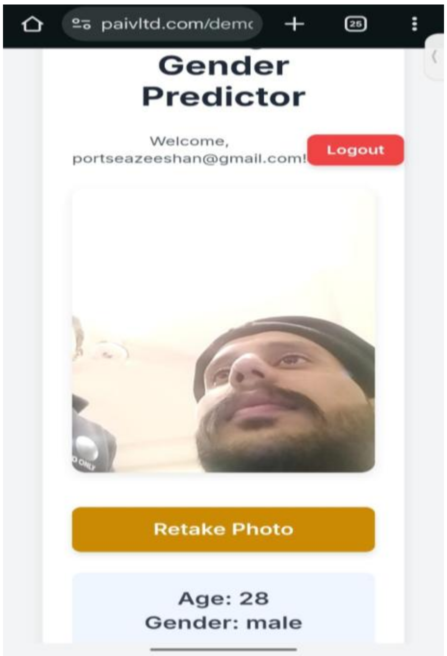

Figure 8: Gender Classification Performance

Figure 9: Gender Classification Performance

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

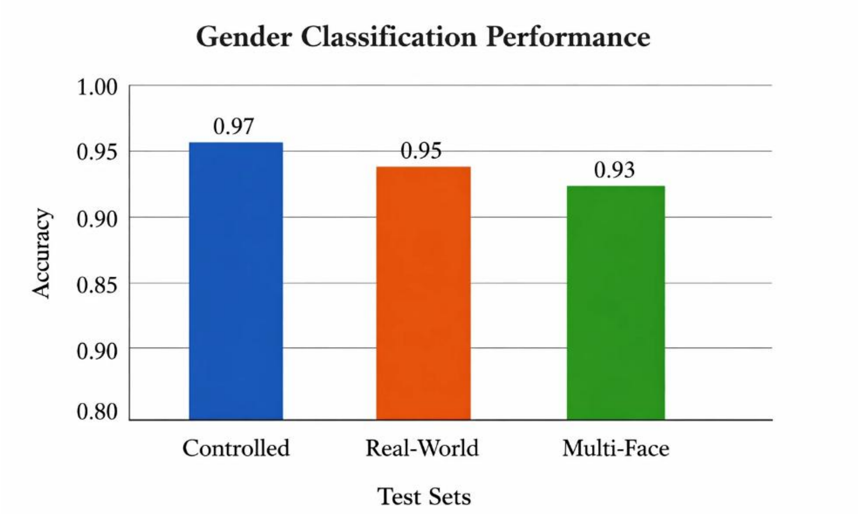

Figure 10. Gender Classification Performance

Age Prediction Accuracy

Another measurement of evaluation of the PAIV system was age prediction. By contrast to gender classification, age estimation has in itself more difficulties, which are expected by the inherent differences in facial appearance with age, as well as myriad weaknesses to athlete techniques like lighting, face expression, and head position. In response, PAIV age estimates (by predefined age groups e.g., 15-20, 21-25, 26-30 years) instead of age regression which is more vulnerable to error. Mean absolute error (MAE) was used to measure system performance and this is the measure of the average difference between the age that has been predicted and the real age. In all test sets, the average MAE of PAIV was five years meaning that in most cases the predicted age fell within the range of five years of the actual age.

This degree of precision is objective considering that the problem of estimating age using real-time video streams is complicated by uncontrolled conditions. Besides MAE, the system was evaluated in terms of assigning individuals to right age group. The age group classification accuracy of PAIV was 87 with high results in the younger age brackets (1520 and 2125 years) which showed higher consistency of the facial features. Though there was a slight reduction in accuracy with age groups above 40 years, still the system showed good classification performance in all age groups. These findings demonstrate the power of CNN-based age prediction model that is able to learn age-related facial pattern and generalize over other differences in terms of ethnicity, light, and face expression. The low prediction error and the high accuracy of age-groups indicate that PAIV can be used in the area of age verification, access control to age-restricted content, and personal content recommendation.

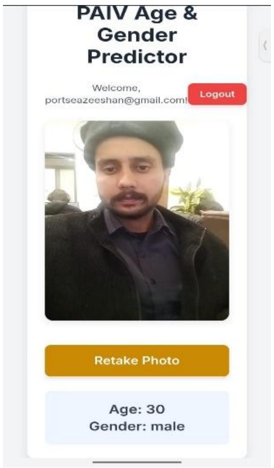

Figure 11: Age Prediction Accuracy

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

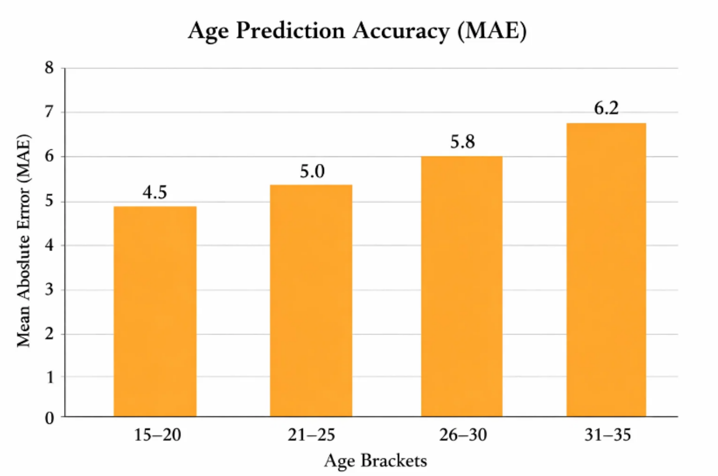

Figure 12. Age Prediction Accuracy (MAE)

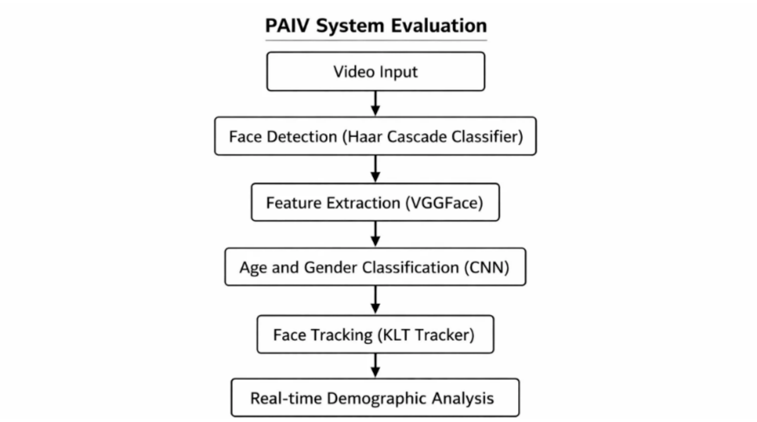

Figure 13: Age Group Classification Accuracy.

Real-Time Processing Performance

The main design requirement of the PAIV system was low-latency processing to guarantee the proper execution within the real-time aspect. The performance of the system under regular (high-resolution: 1080p) and low-resolution (480p) video streams was measured to determine the capability of the system to satisfy the real-time processing demands under different conditions. The desired frame rate was 30 frames per second (FPS), which is vital in applications like live streaming, content moderation and video conferencing. The findings of the experiment prove that PAIV satisfies the real-time performance requirements.

In the case of high-resolution video, the system recorded an average of about 33 ms/frame processing time, which is 30 FPS and showed no performance reduction or notable variation in real-time performance. In low-resolution video, processing speed was further improved to 50 FPS with an average frame processing time of 20 ms, which showed the possibility of good scalability in the context of inputs in terms of resolution. These findings indicate that PAIV exhibits low-latency through a variety of video resolutions and can adjust well to a variety of computational requirements. Video streaming capability offers real-time video processing, which suits applications where a demographic analysis needs to be conducted immediately, such as live streaming video conferencing and automated content moderation.

System Scalability

Scalability is paramount to real-time systems, which are currently implemented on large-scale platforms like companies like social media, e-commerce, and live streaming services. In order to test scalability, PAIV was experimented with in a multi-face and a multi-stream video to determine the behaviour of the system under higher load to maintain accuracy and processing velocity. In the multi-face evaluations, the system was selected on a video

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

frame that had a maximum of 10 faces at a time. PAIV was effective in identifying, tracking and classifying the age and gender of every person in an extremely accurate manner, even when there are partial occlusions and overlapping faces. These findings indicate the strength of the system even in a highly visualized environment. Multi-stream PAIV was a multi-stream video evaluator that used several video streams in real time. The system exhibited both steady throughput and low latency under the assumption of classification accuracy without performance degradation, as well as exhibited high performance in high-demand applications like live streaming platforms, where analysis of large populations of users is needed in terms of demographics. In general, the scalability experiments prove the fact that PAIV can effectively serve with a higher face density in frames and the number of simultaneously processed video streams. Its modular and scale architecture provides the ability to be deployed in applications of a wide variety of applications, from small application scenarios to large-scale enterprise-level applications that demand real-time analysis of an entire video feed.

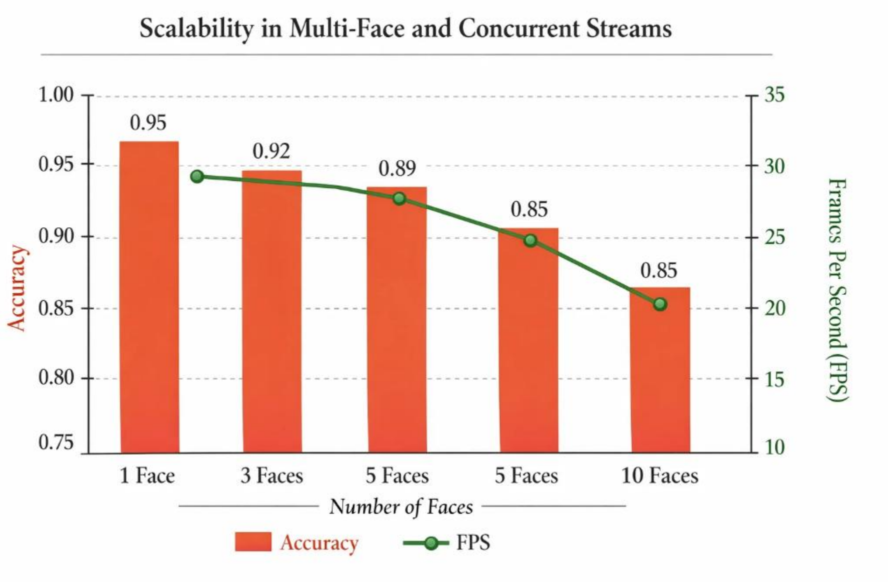

Figure 14: Scalability in Multi-Face and Concurrent Streams

Moreover, the optical scheme of the face tracking in the frames of the video provided with the Kanade-Lucas-Tomasi (KLT) tracker is another innovation that will differentiate PAIV from other systems. Although most of the face recognition systems use repetition of face detection within a frame, the capability of PAIV to track faces across multiple frames minimizes the computational costs and enhances efficiency.

DISCUSSION

The PAIV system was found to be good in real-time age and gender classification on video streams with accuracy 95 percent in gender classification and the Mean Absolute Error (MAE) of video age prediction with a result of ±5. The results suggest the effectiveness of learning deep learning models with real-time face tracking to demographic analysis, and they can be deployed to several applications, including surveillance, user profiling, and content leveraging. The use of multi-output Convolutional Neural Networks (CNNs) and VGGFace and real-time tracking of the faces allowed the system to achieve extremely high accuracy rate despite the difficulty in handling real-time video images. It is also very impressive that it works in real time and has a rate of 30 frames per second (fps) considering the computational and processing power required to face the video-based demographic analysis [9]. Live streaming applications will require live classification since the demographic data of the users is to be processed in real-time [10]. The PAIV system can also claim how deep learning models can be adapted to work in real-time without affecting accuracy by guaranteeing that face tracking and detection are performed during the entire video. This is what makes the PAIV system different when compared to most of the traditional systems that find it hard to keep the balance of speed and accuracy in a video based system.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Nevertheless, even though the accuracy of the genders classification and the age estimation made by the PAIV system are impressive, multiple limitations also require a solution [11]. The mean absolute error (MAE) of age prediction is one of the main obstacles with the value of ±5%. This error margin can be considered competitive with other currently available estimate techniques, though age estimation is also not an easy task since facial features vary among various age groups.

The fact that the system is based on fixed ranges of age also contributes to the lack of accuracy in the system and determining the actual ages is still a persistent problem of the research on facial recognition [12]. Past research has discovered that even the state-of-the-art models like deep neural networks or transfer learning models have challenges when differentiating among the age groups with similar facial characteristics like in the adulthood age range (30-40 years). Nevertheless, the predictive age capability of the PAIV system even in the age ranges may find useful application in other practical tasks, where age precision may not be require [13].

Other drawbacks of PAIV system include its performance in weak light conditions and low resolution of video input. Although the system works in confined situations, the real world is usually affected by low-quality videos caused by different factors such as low camera quality, low light, or camera shaking [14]. Such problems may affect face recognition and, consequently, precision in terms of age and gender recognition. A great number of the existing approaches to the issue of computer vision have the same challenge, and the enhancement of the system to such variations still is one of the main directions of the further work. This limitation could be overcome through techniques like super-resolution imaging, noise reduction filters and better pre-processing.

Moreover, the PAIV system now employs a fairly primitive face detection methodology relying upon Haar Cascade Classifiers that are computationally inexpensive yet that might lack capability in operating in harsh environments such as extreme pose change, or partial cover [12]. Deep learning-based detectors may be included to make the face detectors more precise and operational under a variety of conditions, including Single Shot Multibox Detector (SSD) or Faster R-CNN. These detectors have demonstrated high performance of face detection under different circumstances such as low-light and extreme face positions.

Another challenge encountered by the system is that of detecting multiple faces in a single frame which is a common situation in video streams. Although it has been shown that the KLT tracker can be used well in tracking of one face per frame, further processing power and even delay when handling multiple faces is being introduced which in turn can lead to delays especially in real-time applications. Multi-face tracking is a current field of operation and techniques, including multi-task learning which simultaneously detects and classifies many faces, may greatly enhance the capability of the PAIV system to deal with multi-faces without compromising the speed [15].

As to future state of improvement, various directions are possible to explore the performance of the PAIV system. To begin with, the age prediction model used should be enhanced by shifting to continuous age regression rather than relying on the age ranges so as to have an accurate age prediction [6]. It would mean that the CNN model would need to be optimized to reproduce specific values in terms of age and the authors could use larger and more diverse data volumes in order to embrace the subtle difference between different age groups.

Also, the inclusion of time aspect, such as the variation of the face features with time or age pattern of facial expression, would enhance the accuracy. The next possible improvement is the incorporation of the multi-modes data whereby the system may embrace a three-dimensional data comprising of facial recognition of other biometric indicators such as voice or gestures, to give a more precise demographic profile [2].

In addition, the weaknesses of dealing with the dark setting and low-resolution videos cannot be ignored as a priority. In the second stage of age and gender estimation, it would also be possible to apply advanced image processing methods to enhance the quality of input frames, including image denoising or super-resolution networks [16]. These techniques would improve the system's performance in real-world scenarios where video customer demands can vary greatly. Finally, expanding the system's ability to work with multiple faces per frame is crucial to improving its practical usability. More advanced multi-face recognition techniques, such as the addition of multi-object tracking algorithms, could enhance future PAIV system implementations by enabling the system to accurately identify several people in real time.

Journal of Information Systems Engineering and Management
2026, 11(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The PAIV system is an important advancement in real-time categorisation of videos according to gender and age. Even though the system has significant potential of classifying gender and approximating age, it can still be enhanced on accuracy and tolerance to the fluctuations in various conditions. (Savchenko, 2019). With the assurance of the said limitations the PAIV system will be capable of bringing further innovation to the automated video-based demographic analysis and serve as a more dependable tool in a real-life situation (surveillance, content moderation, and customized experiences in the media).

REFRENCES

[1] D. Ali, F. Epifania, N. Ahmed, B. Khan, L. Marconi, and R. Matamoros Aragon, "A modified CNN for age and gender prediction," in *CEUR Workshop Proceedings*, 2022, vol. 3463: CEUR-WS, pp. 1-7.

[2] K. Bešenić, J. Ahlberg, and I. Pandžić, "Let Me Take a Better Look: Towards Video-Based Age Estimation," in *13th International Conference on Pattern Recognition Applications and Methods-ICPRAM*, 2024, pp. 57-59.

[3] S. T. Challa, S. Jindam, R. R. Reddy, and K. Uthej, "Age and gender prediction using face recognition," *International Journal of Engineering and Advanced Technology (IJEAT),* vol. 11, no. 2, pp. 48-51, 2021.

[4] H. Du, H. Shi, D. Zeng, X.-P. Zhang, and T. Mei, "The elements of end-to-end deep face recognition: A survey of recent advances," *ACM computing surveys (CSUR),* vol. 54, no. 10s, pp. 1-42, 2022.

[5] S. Haseena, S. Saroja, R. Madavan, A. Karthick, B. Pant, and M. Kifetew, "Prediction of the age and gender based on human face images based on deep learning algorithm," *Computational and mathematical methods in medicine,* vol. 2022, no. 1, p. 1413597, 2022.

[6] M. Koconń and S. Pawlukiewicz, "Age Estimation and Gender Classification from Facial Images," *Applied Sciences,* vol. 15, no. 18, p. 10212, 2025.

[7] N. Lakshmiprabha, "Face image analysis using aam, gabor, lbp and wd features for gender, age, expression and ethnicity classification," *arXiv preprint arXiv:1604.01684,* 2016.

[8] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 34-42.

[9] O. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *BMVC 2015-Proceedings of the British Machine Vision Conference 2015*, 2015: British Machine Vision Association.

[10] W. Pei, H. Dibeklioğlu, T. Baltrušaitis, and D. M. Tax, "Attended end-to-end architecture for age estimation from facial expression videos," *IEEE Transactions on Image Processing,* vol. 29, pp. 1972-1984, 2019.

[11] A. V. Savchenko, "Efficient facial representations for age, gender and identity recognition in organizing photo albums using multi-output ConvNet," *PeerJ Computer Science,* vol. 5, p. e197, 2019.

[12] A. Subuhi, J. Sathwik, and S. Kumar, "CNN-Based Facial Analysis For Precise Age and Gender Recognition," in *International Conference on Computer Science and Communication Engineering (ICCSCE 2025)*, 2025: Atlantis Press, pp. 1496-1507.

[13] A. Tunc, S. Tasdemir, M. Koklu, and A. C. Cinar, "Age group and gender classification using convolutional neural networks with a fuzzy logic-based filter method for noise reduction," *Journal of Intelligent & Fuzzy Systems,* vol. 42, no. 1, pp. 491-501, 2021.

[14] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, 2001, vol. 1: Ieee, pp. I-I.

[15] A. Nazare and S. Padmannavar, "An Integrated Approach for Real-Time Gender and Age Classification in Video Inputs Using FaceNet and Deep Learning Techniques," *International Journal of Advanced Computer Science & Applications,* vol. 15, no. 7, 2024.

[16] B. Abirami, T. Subashini, and V. Mahavaishnavi, "Gender and age prediction from real time facial images using CNN," *Materials Today: Proceedings,* vol. 33, pp. 4708-4712, 2020.