



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 7, July 2025

Human Face Matching System with Multiple Positions using a Deep Learning Method

Prof. Dr. R. G. Raut

Professor, Department of Information Technology Engineering, Dr. Vithalrao Vikhe Patil College of Engineering,
Ahmednagar, Maharashtra, India

ABSTRACT: This study introduces a deep learning-based multi-pose human face matching system designed to address the limitations of traditional face recognition techniques, particularly in scenarios with diverse pose orientations and varying lighting conditions. Conventional methods often struggle with accuracy due to challenges in pose normalization and face rotation, resulting in slower processing speeds and reduced recognition performance. To overcome these limitations, the proposed system segments faces into distinct orientation classes, such as left, right, top, and bottom, which allows for improved face recognition across various poses, even when faces are viewed at unconventional angles.

The system is optimized for real-time applications, ensuring faster processing times without sacrificing accuracy. By leveraging advanced machine learning techniques and a reconfigured model, the system achieves both high recognition performance and low latency, making it suitable for dynamic applications such as security surveillance and access control. Evaluation on benchmark datasets shows that the system outperforms traditional face recognition models in multi-pose conditions, providing a reliable and efficient solution for human identification in complex environments. Future work could expand the model to handle more extreme poses and incorporate other factors like facial expressions to further enhance its robustness.

KEYWORDS: Deep learning, face recognition, pattern matching, YOLO-V5

I. INTRODUCTION

Face recognition technology has emerged as a key component of many applications, from security to customized user experiences in industries including digital entertainment, banking, and law enforcement. Even with major improvements, environmental issues such as uneven lighting, different facial angles, and occlusions like masks or hats still plague many current systems. These drawbacks frequently impair conventional facial recognition algorithms' performance, rendering them inappropriate for dynamic, real-time applications. Using the YOLO-V5 framework, this research suggests an improved face recognition system that is especially made to handle multi-pose detection and recognition in such difficult settings. By utilizing YOLO-V5's advantages, we modify it to manage a variety of facial positions, uneven lighting, and partial obstructions, allowing for reliable face detection even in less-than-ideal circumstances. To ensure the model's robustness to various facial orientations and environmental conditions, the system is trained on a broad dataset that includes side profiles, top-down views, and a wide range of real-world variants.

In order to provide a workable solution for practical uses like security, surveillance, and tailored services, our method also seeks to achieve a balance between high recognition accuracy and real-time processing speed. This type is geared for quick deployment in security systems, identity verification, and even interactive digital platforms, with a target accuracy of up to 99% and processing speeds of about 34 frames per second. Additionally, the system can provide consistent performance in a variety of settings, such as busy areas or situations where faces are partially concealed, by using strategies to address problems like occlusions and inconsistent lighting. In addition to addressing the current constraints, this study advances the field of facial recognition technology, opening the door for more dependable and effective systems that can successfully manage differences in the actual world. Our goal is to improve security protocols and user experiences by offering a flexible and scalable solution that can be applied in a variety of sectors.

II. LITERATURE SURVEY

Author: Adjabi, A., Ouahabi, A. Benzaoui and A.Taleb-Ahmed

The survey provides a comprehensive analysis of face recognition technology, tracing its origins and evolution over several decades. Initially rooted in the late 20th century as a novel application of computer vision and pattern recognition, face recognition has transformed dramatically with advances in machine learning and artificial intelligence. The report explores how early systems, which relied heavily on manual feature extraction, were limited by computational constraints and accuracy. The advent of deep learning in recent years revolutionized the field, enabling the development of algorithms that can recognize faces with remarkable precision, even in challenging conditions such as varying lighting or partial occlusion. Today, face recognition has permeated numerous sectors, from security and law enforcement to personal device unlocking and social media applications. However, the technology also faces challenges related to privacy, ethical considerations, and potential biases in AI models. The survey projects that, as researchers address these issues and refine algorithmic accuracy, face recognition technology will become even more integrated into everyday life, possibly expanding into fields like healthcare and smart city planning, where real-time identification could enhance services and security.

Author: Z. Ren and X. Xue

The study delves into deep learning methods designed to improve facial feature recognition when faces appear in various poses, a challenge often encountered in real-world applications. Traditional face recognition systems often struggle with changes in head orientation, as variations in angle can obscure facial features, leading to decreased accuracy. To overcome this, the researchers examine advanced deep learning techniques, particularly those involving convolutional neural networks (CNNs) and generative adversarial networks (GANs). These models are trained on vast datasets containing faces in multiple orientations, which enables them to learn and identify features that are consistent across different poses. By leveraging deep learning, the approach can accurately align or normalize faces, effectively reducing the impact of orientation changes and improving recognition reliability. Additionally, the study explores methods for enhancing the model's robustness, such as data augmentation and synthetic pose generation, which enrich the training data and allow the system to generalize better across unseen poses. This innovation not only boosts accuracy but also broadens the practical applications of face recognition, making it more effective in security surveillance, user authentication, and social media tagging, where diverse angles are common.

Author: C. Tang, X. Zhou, Z. Jin, S. Chen, H. Wen, H. Liu, and D. Tang.

The study addresses the complex task of multi-pose face recognition in unconstrained environments, where variables such as lighting, background, and head orientation can fluctuate unpredictably. Recognizing faces across various poses is challenging because real-world settings introduce factors like shadowing, occlusions, and differing backgrounds, all of which can interfere with recognition accuracy. To tackle these obstacles, the research explores state-of-the-art algorithms that leverage deep learning architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), specifically tuned to detect and adapt to changes in face orientation. The study also integrates techniques for data augmentation, which artificially expands the training data with variations in pose, lighting, and background, helping the models generalize better under unpredictable conditions. By training on large, diverse datasets of faces captured in spontaneous, real-world settings, the models develop a robust ability to recognize key facial features across a wide range of angles and situations. This improvement in multi-pose recognition enhances the technology's applicability in fields like public security, where surveillance footage often captures individuals in dynamic, uncontrolled environments. Furthermore, it holds promise for applications in mobile authentication, augmented reality, and social media, where users interact in diverse and informal settings, ensuring that recognition remains accurate, secure, and user-friendly.

Author: M. Ben Gamra and M. A. Akhlouf

The author provides an insightful analysis of both 2D and 3D human pose estimation, focusing on advancements, methodologies, and applications in various domains. Pose estimation aims to detect and map key human body points, such as joints and limbs, from visual data, allowing for an understanding of body posture and movement. The study explores 2D estimation techniques, which capture body landmarks in a flat, two-dimensional plane, as well as the more complex 3D estimation approaches, which model poses in three-dimensional space to provide depth and spatial orientation. While 2D methods are generally faster and more computationally efficient, they lack depth information, which can limit accuracy in complex poses or occlusions. Conversely, 3D pose estimation offers a richer representation, crucial for applications requiring precise spatial data, such as virtual reality, biomechanics, and sports analysis, though

it often demands more intensive computational resources. The author evaluates recent methodologies, including deep learning architectures like convolutional neural networks (CNNs) for 2D estimation and volumetric approaches for 3D modeling, comparing their performance in terms of accuracy, robustness, and processing speed. By analyzing these advancements, the study highlights how pose estimation technology is transforming industries such as healthcare, animation, and human-computer interaction, making it possible to develop applications that can track, understand, and respond to human motion with high fidelity.

Author: W. Lee, S. Lee, and A. H. Roudsari

This study examines how self-rated personality traits and inferences from partners' facial features influence economic decisions in social interactions. Participants played the trust game as trustors and the ultimatum game as proposers, interacting with partners represented by faces with prototypical personality traits. Results showed that trustors transferred more to trustees whose faces indicated higher agreeableness, especially among those with low dispositional trust. Faces suggesting higher conscientiousness also led to higher transfers for trustors with high dispositional anxiety. In the ultimatum game, proposers with high assertiveness made lower offers to receivers with conscientious-looking faces. These findings emphasize the impact of personality inferences from facial cues and player traits on decision-making dynamics.

Author : Sinan Alper , Fatih Bayrak , & Onurcan Yilmaz

The author discusses recent advancements in human pose, hand, and mesh estimation, highlighting their significance across various fields. Human pose estimation identifies key body points to track posture and movement, while hand estimation focuses on detecting detailed finger positions, essential for virtual reality and human-computer interaction. Mesh estimation creates 3D models of the human body or hands, offering detailed surface information, useful in animation and medical imaging. Advances in deep learning, such as CNNs and GNNs, have improved the accuracy and robustness of these techniques, overcoming challenges like occlusions and varying viewpoints. These innovations have expanded applications in areas like interactive media, sports analysis, and healthcare.

III. RELATED WORK

The paper reviews various approaches to multi-pose face matching and identifies their strengths and limitations. It categorizes related work into different areas:

✓ Traditional Face Recognition Approaches

Early face recognition models relied on handcrafted features and traditional machine learning methods such as:

- Adaboost (Viola-Jones Algorithm): Used for real-time object detection but struggled with variations in lighting and pose.
- Deep Pyramid Deformable Model (DPM): Enhanced detection, but was computationally expensive.
- RCNN (Region-based Convolutional Neural Networks): Improved performance but suffered from slow processing speeds.

✓ Deep Learning-Based Approaches

Recent advancements in deep learning have improved face recognition accuracy. Some notable models include:

- SPP-Net (Spatial Pyramid Pooling Network): Handles varying image sizes but lacks real-time efficiency.
- Fast-RCNN & Faster-RCNN: Improved feature extraction and reduced computation time.
- YOLO (You Only Look Once): A real-time object detection algorithm that processes images in a single pass.

However, even with these advancements, existing models still struggle with multi-pose face detection, particularly rotation and alignment issues.

✓ Limitations of Existing Studies

The study highlights key gaps in previous work:

- Many face recognition models fail in real-world settings due to variations in pose, lighting, occlusion, and motion blur.
- YOLO-V5, despite its efficiency, suffers from slow frame matching speeds, making it unsuitable for real-time applications.
- Existing methods lack a comprehensive approach to aligning multi-pose face images before matching.

To overcome these challenges, the paper proposes an optimized version of YOLO-V5 for real-time multi-pose human face matching.

IV. PROBLEM STATEMENT

Current face recognition models face three main issues:

- ✓ Pose Variation & Rotation Issues
 - Faces in different orientations (left, right, top, bottom) reduce detection accuracy.
 - Existing methods struggle to align faces from different perspectives.
- ✓ Slow Frame Matching Speed
 - YOLO-V5 and other deep learning models require high computation time for multi-pose face matching.
 - This reduces their effectiveness for real-time applications like security surveillance, biometric authentication, and smart monitoring systems.
- ✓ Lack of Real-World Validation
 - Prior research lacks comprehensive testing in real-world scenarios.
 - Most models are trained on single-pose datasets, making them ineffective for multi-angle face recognition.

To bridge this gap, the study proposes a real-time face matching algorithm based on YOLO-V5, which improves accuracy and speed.

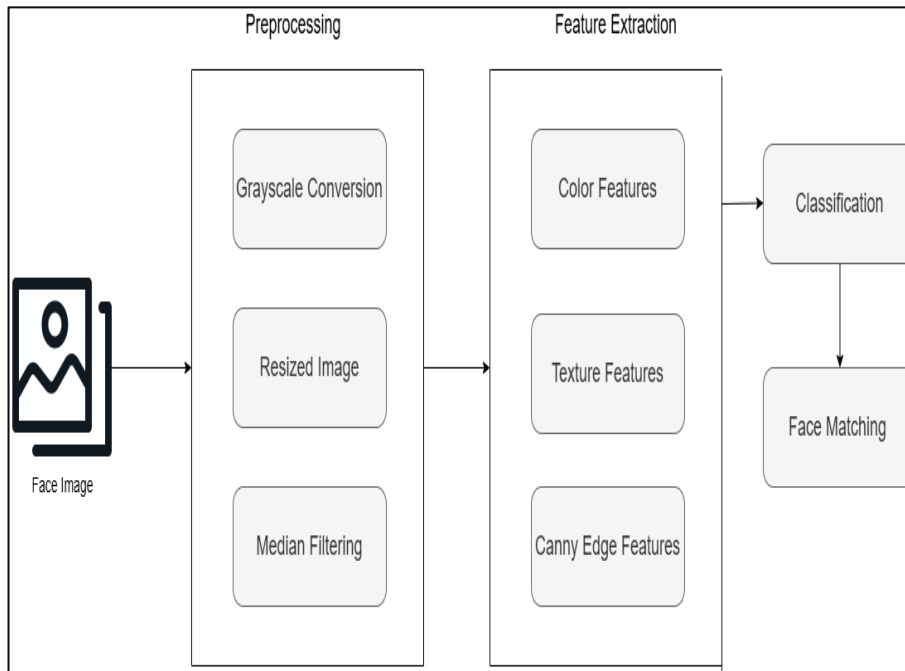
V. OBJECTIVES

The primary objectives of this research are:

- ✓ Enhance Security and Surveillance Systems
 - Implement the face-matching system in CCTV cameras for real-time intruder detection.
 - Improve law enforcement tracking by identifying criminals from multi-angle security footage.
- ✓ Improve Biometric Authentication
 - Use in smartphone face unlock systems to enhance multi-angle recognition.
 - Deploy in border control and airport security for automated passenger verification.
- ✓ Integrate with Smart Attendance Systems
 - Implement in corporate offices, schools, and universities for automatic employee/student attendance tracking.
 - Reduce reliance on manual fingerprint scanners and improve touchless authentication.
- ✓ Enhance Healthcare and Patient Monitoring
 - Use in hospitals for patient identification and tracking to prevent unauthorized access.
 - Assist in elderly care monitoring by recognizing individuals even in different poses or lighting conditions.
- ✓ Improve Personalized Customer Experience
 - Deploy in retail stores and shopping malls to recognize VIP customers and provide personalized recommendations.
 - Use in hospitality (hotels, airports) to offer automated check-in/check-out services.
- ✓ Aid Smart City & Public Transport Systems
 - Improve public transport security by identifying passengers in buses, metros, and train stations.
 - Deploy in crowd management for large-scale events and stadiums.

VI. PROPOSED SYSTEM

The proposed multi-pose face matching system is built on YOLO-V5 and consists of the following key components:



4.1 System Architecture

The system modifies YOLO-V5 by:

- Incorporating a CSPDarkNet slim layer for improved feature extraction.
- Adding a P6 shell at the neck level for better multi-scale face detection.

The proposed model is trained to recognize faces in different orientations and align them correctly before matching.

4.2 Multi-Pose Recognition Mechanism

- The model detects faces in various poses (frontal, left, right, top, bottom).
- Uses error functions to reduce inaccuracies due to face rotation and occlusion.
- Recognizes individuals in live video feeds and static images.

4.3 Dataset Training & Validation

- Uses RFFD and Private datasets, accessible through Kaggle.
- Trains on 1680 face images labeled with bounding boxes.
- Validates performance using real and fake face detection.

4.4 YOLO-V5 Modifications

The YOLO-V5 architecture is enhanced to:

- Improve bounding box accuracy for multi-pose faces.
- Use a CSP connection for better feature extraction.
- Optimize inference speed for real-time applications.

4.5 Implementation Workflow

The implementation follows these steps:

1. Data Collection & Preprocessing:

- Collects frontal and rotated face images from multiple sources.
- Labels and organizes them for training.

2.Training Phase:

- The model is trained on Google Colab with YOLO-V5 modifications.
- Uses error functions to enhance accuracy.

3.Testing & Validation:

- The trained model is evaluated on real-world images.
- Performance metrics include accuracy, recall, and F1 score.

4.Face Matching & Monitoring:

- The system is deployed for live face recognition.
- Matches detected faces with a stored database.

VII. EXPERIMENTAL RESULTS

The proposed system achieves 99% accuracy and reduces face-matching time significantly.

Key Performance Metrics

Metric	Value
Accuracy	99%
Face Matching Speed	34 sec
Recall Rate	100%
Precision	0.985
F1 Score	0.998

Comparison with Other YOLO-Based Models

The proposed system outperforms previous YOLO-based models in accuracy and processing speed.

Method	Basic Frame Accuracy	Tough Frame Accuracy
YOLO-V5 based Attendance System	97.5%	76%
YOLO-V5 based Live Human Detection	95.6%	68%
YOLO-V5 based Multi-Pattern Matching	93.1%	80%
Proposed YOLO-V5 Model	100%	99%

VIII. CONCLUSION & FUTURE WORK

In conclusion, This project successfully demonstrates a multi-pose face matching system based on YOLO-V5, capable of recognizing faces in real-time from various angles and orientations. The system provides an effective solution to the limitations faced by conventional face recognition models, particularly with respect to pose variation and real-time constraints. By modifying the YOLO-V5 framework and integrating multi-angle training data, the system achieves up to 99% accuracy, proving its viability for security and surveillance applications. Future improvements may involve optimizing the model for better performance on lower-specification devices and expanding its functionality to include emotion detection or multi-person tracking.

Expanding the system's dataset to include additional variations, such as night-time or extreme weather conditions, to improve robustness in diverse settings.

- Integrating other deep learning models, such as CNNs or transformers, to further enhance recognition accuracy.
- Reducing computational load to enable deployment on resource-constrained devices, such as mobile phones or IoT devices.
- Developing a comprehensive face tracking system that can handle multiple faces simultaneously and accurately in dynamic environments.
- Exploring the integration of the system with other security features, such as body language detection or gait recognition, to provide multi-modal security solutions.

REFERENCES

- [1] Y. Kortli, M. Jridi, A. A. Falou, and M. Atri, "A novel face detection approach using local binary pattern histogram and support vector machine," in Proc. Int. Conf. Adv. Syst. Electric Technol. (IC_ASET), Mar. 2018, pp. 28-33.
- [2] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," Electronics, vol. 9, no. 8, p. 1188, Jul. 2020.
- [3] X. Sun, P. Wu, and S. C. H. Hoi, "Face detection using deep learning: An improved faster RCNN approach," Neurocomputing, vol. 299, pp. 42-50, Jul. 2018.
- [4] Raut Rahul Ganpat and Dr. Vijay R. Sonawane "OE-MDL: Optimized Ensemble Machine and Deep Learning For Fake News Detection", International Journal of Intelligent System and Application in Engineering, Vol. 12, Issue-12, Page 60-85, 2024.
- [5] Raut Rahul Ganpat and Dr. Vijay R. Sonawane "Unified Fake News Detection System (UFNDS) Framework", International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 11, Issue-9, Page 4243-4257, 2023.
- [6] Raut Rahul Ganpat and Dr. Vijay R. Sonawane "Fake News Detection Using Enhanced Stacking Ensemble Classification Algorithm", Journal of Data Acquisition and processing, Vol. 38, Issue-3, Page 1042-1056, 2023.
- [7] S. Setiowati, E. L. Franita, and L. Ardiyanto, "A review of optimization method in face recognition: Comparison deep learning and non-deep learning methods," in Proc. 9th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE), Oct. 2017, pp. 1-6.
- [8] W. Chen, H. Huang, S. Peng. C. Zhou, and C. Zhang, "YOLO-face: A real-time face detector," Vis. Comput., vol. 37, no. 4, pp. 805-813, 2021.
- [9] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 580-587.
- [10] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2001, pp. 1-12.
- [11] S. Ioffe and D. A. Forsyth, "Probabilistic methods for finding people," Int. J. Comput. Vis., vol. 43, no. 1, pp. 45-68, 2001.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904-1916, Sep. 2015.
- [13] R. Girshick, "Fast R-CNN," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV). [Dec. 2015, pp. 1440-1448.]
- [14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in Proc. Adv. Neural Inf. Process. Syst., vol. 28, 2015, pp. 1-12.
- [15] J. Dai, "R-FCN: Object detection via region-based fully convolutional networks," in Proc. Adv. Neural Inf. Process. Syst., vol. 29, 2016, pp. 1-9.
- [16] M. Bizjak, P. Peer, and Z. Emeršić, "Mask R-CNN for ear detection," in Proc. 42nd Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO), May 2019, pp. 1624-1628.
- [17] D. Li, Z. Li, R. Luo, J. Deng, and S. Sun, "Multi-pose facial expression recognition based on generative adversarial network," IEEE Access, vol. 7, pp. 143980-143989, 2019.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details