International Scientific Journal of Engineering and Management
Volume: 03 Issue: 05 | May – 2024
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

YOLO-FaceNet Fusion: Innovative System for Facial Recognition and Feature Extraction

Mr. Mukesh Kumar[1], Aman Kumar[2], Ankush Gupta[3], Gaurav Kapoor[4], Vineet Vashishtha [5]

Assistant Professor[1], student[1], student[2], student[3], student[4]

(Department of Computer Science and Engineering: Internet of Things)

Meerut Institute of Engineering and Technology, Meerut

mukesh.cs@miet.ac.in[1], aman.kumar.cseiot.2020@miet.ac.in[2], ankush.gupta.cseiot.2020@miet.ac.in[3], gaurav.kapoor.cseiot.2020@miet.ac.in[4], vineet.vashishtha.cseiot.2020@miet.ac.in[5]

Abstract:

Facial verification is vital in domains like Heathcare, Identity verification, etc. study presents a sophisticated facial verifying system utilizing state-of-the-art technologies: ArcFace is utilized for the edification of the model, while FaceNet is employed for the elicitation of facial characteristics. Initially, YOLO is deployed to discern and extricate visages from the input effigies. Subsequently, FaceNet is utilized to derive high-caliber facial attributes. In the final stage, ArcFace is employed for the model's edification, thus augmenting its robustness and precision. The apparatus gauges facial resemblance employing the Euclidean distance metric and is appraised through measures such as recall, accuracy, and precision. This approach yields a dependable and accurate facial verification system, thereby enhancing operational efficiency in practical applications..

Keywords – You Only Look Once (YOLO), ArcFace, Convolutional Neural Network (CNN), FaceNet, Machine Learning, Deep Learning, Euclidean Distance**.**

1. Introduction:

Facial verification, which involves determining if two face images correspond to the same individual, is vital in numerous applications, such as digital identity authentication and security systems. Recent advancements in deep learning have significantly improved face verification

International Scientific Journal of Engineering and Management ISSN: 2583-6129
Volume: 03 Issue: 05 | May – 2024 DOI: 10.55041/ISJEM01802
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

systems. This paper presents a system combining the Euclidean distance measure, FaceNet face recognition, and YOLO object detection for excellent face verification accuracy [1] [2].

The proposed system initiates by employing YOLO object detection to swiftly and accurately identify faces in input images. Following detection, the identified faces are cropped and aligned to minimize variations in lighting and positioning [3] [4].

The Euclidean distance metric is used to evaluate the similarity between the feature vectors of two facial images. Model training with the ArcFace loss function enhances the accuracy of the facial recognition model by optimizing angular margins between features belonging to distinct classes [5] [6].

The accuracy of the facial recognition model has been increased with the application of the ArcFace loss function.[7] [8].

2. Literature Review

A plethora of inquiries have been undertaken, with certain ones acting as the impetus for this exploration. Here are brief summaries of some of these important, recent works:

Insaf Adjabi et al.'s study [1] contrasts 2D and 3D facial recognition techniques. They discover that whereas 2D techniques perform best in controlled settings, 3D strategies have an edge when handling different lighting conditions and points of view.An alternate method, the 3D approach, was proposed. 3D data is more effective for recognition algorithms since it is not affected by position or illumination. This study delves into the chronicles of facial recognition technology, exploring the latest avant-garde methodologies and contemplating prospective advancements.

Yassin Kortli et al. [2] offer a spectrum of facial recognition methodologies, each possessing distinctive merits and demerits concerning resilience, precision, and intricacy. These methodologies encompass local, holistic, and amalgamated approaches. They offer a comprehensive evaluation of different methods, demonstrating how well they perform on widely used datasets for both supervised and unsupervised learning. Along with discussing the experimental circumstances and issues that these techniques address, the report presents numerical results for the most promising approaches.

ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

Murat Taskiran et al. [3] compiled procedures for gathering and classifying facial biometric data, compiled by who also list methods for classifying face recognition into image-based and video-based categories. They examine notable datasets used in face identification research, highlight historical developments, and emphasize important processing steps. Moreover, they analyze current facial recognition methodologies built upon deep learning and offer suggestions for future research directions.

Ashu Kumar et al.'s study [4] furnishes a comprehensive disquisition on the myriad techniques scrutinized for visage recognition in digital images. This treatise further examines various applications and quandaries pertinent to face detection. Moreover, the attributes of numerous renowned face detection repositories are delineated. They organize specific talks on the practical issues to develop a trustworthy face detection system. They provide us with a few interesting research topics for our next reading.

Iacopo Masi et al. [5] delineated principal progressions acquiring face representations for identification and verification, along with advancements in deep face recognition. Within the last five years, state-of-the-art (SOTA) face recognition algorithms have started to appear in premier computer vision contexts. The survey explains the principle in an organized and understandable manner. The questionnaire is structured into multiple sections that follow a standard facial recognition workflow: Which public data sets were utilized for SOTA system training? (A) The transfer learning architecture and loss functions; (B) The facial preprocessing element (alignment, detection, etc.); (C) Face detection technology for identification and verification

Anagha S. Dhavalikar et al. [6] proffered a plan for an Automated Visage Expression Recognition System (AVERS). There exist three stages in the recommended method: face detection in (A), feature extraction in (B), and facial emotion identification in (C). The YCbCr color model is used to determine skin color in the initial phase of face detection. Illumination correction is implemented to ensure consistency across the face, and morphological techniques are employed to maintain the pertinent facial area. Utilizing the Active Appearance Model (AAM) method, facial characteristics such as the mouth, nose, and eyes are extracted from the outcome of the initial stage. The third phase, automatic face expression identification, uses the basic Euclidean Distance approach. Using this method, the Euclidean distance between the training and query images' feature points is compared. The minimal Euclidean distance determines the output image expression.

International Scientific Journal of Engineering and Management
Volume: 03 Issue: 05 | May – 2024
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

Madan Lal et al. [7] have offered various techniques for developing recognition systems. Still, the most common operations that involve the iris and fingers. These require user participation or activity in order to access the system. Furthermore, contemporary techniques provide member access without its intervention. For photographs, LFT (Labeled Faces in the Wild) is utilized, while for films, YTM (YouTube Faces) is employed. This ensures that face recognition databases are distinct for controlled images and uncontrolled videos.. A visage identification system comprises three principal constituents: initial processing (A), attribute culling (B), and categorization (C).

Examining the aforementioned studies aids in comprehending the importance of facial recognition systems and their various applications.

3. Proposed Methodology

Gathering face photos of both positive and negative pairings is the first stage in face verification process. Negative pairs consist of pictures of separate people, whereas positive pairs consist of pictures of the same person. Next To eliminate position and lighting changes, Prepare the collected facial images for further processing by cropping and aligning them to standardize their appearance. This stage guarantees that the face photos are appropriate for feature extraction and normalized [9][10]. Now, identify faces in the supplied photographs by using YOLO object detection. Yolo can identify faces and other things in photos with speed and accuracy. Utilizing a pre-trained FaceNet model, extract the features from the detected faces. Each countenance image is thenceforth transmuted into a high-dimensional feature vector by FaceNet, encapsulating the singular characteristics of each visage [11][12].

Upon extracting facial attributes, the resemblance between the feature vectors of two visage images is evaluated using the Euclidean distance metric. Two feature vectors are deemed congruent if their distance falls below a predetermined threshold, signifying that the two facial images pertain to the same individual. Conversely, if the distance surpasses the threshold, they are classified as non-matches, indicating that the two facial images correspond to different individuals [13][14].

The figure below elucidates the Euclidean distance:

International Scientific Journal of Engineering and Management
ISSN: 2583-6129
Volume: 03 Issue: 05 | May – 2024
DOI: 10.55041/ISJEM01802
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

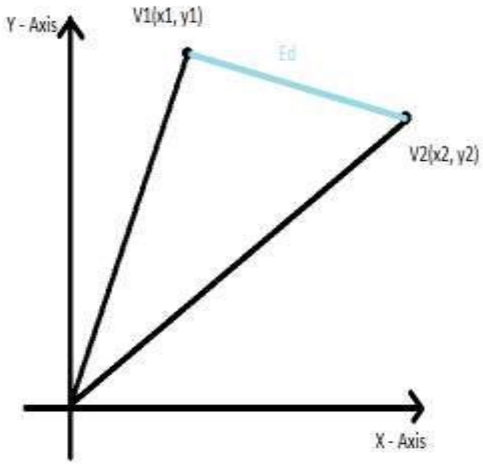

Fig. 1 – Euclidean Distance

The calculating formula is as follows:

$$\sqrt{\sum_{i=1}^{z} (V1i - V2i)^2}$$

Through this training, the network acquires embeddings that exhibit high discrimination, making them more effective for face recognition applications. Since choosing right loss function is crucial to the construction of such a system. Due to the typical softmax loss function's incapacity to account for significant intrafacial fluctuations, ArcFace, also known as Additive Angular Margin Loss, is utilized to distinguish between two facial images. Below is the equation used to compute the ArcFace loss function [15][16][17].

$$L_3 = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{e^{s\left(\cos\left(\theta_{y_i}+m\right)\right)}}{e^{s\left(\cos\left(\theta_{y_i}+m\right)\right)} + \sum_{j=1,j\neq y_i}^{n} e^{s\cos\theta_j}}$$

Finally, assess face verification system's performance using industry-standard measures such as F1 score, accuracy, precision, and recall [18][19].

International Scientific Journal of Engineering and Management
Volume: 03 Issue: 05 | May – 2024
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

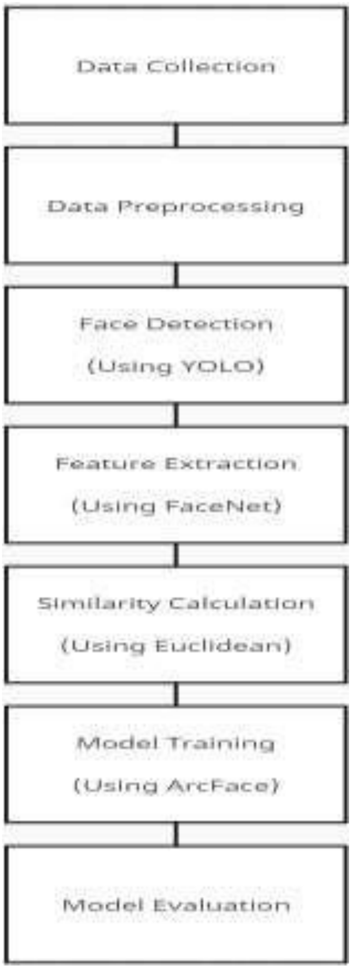

Fig. 2- Data Flow Diagram

Following an examination of the mentioned parameters, the services yield their respective results. [20]. The functionalities provided by this system include:

- **Visage detection:** It is employed to discern all countenances present in the image**.** Its function is limited to detecting the presence and positioning of faces within the image, without the capability to recognize individual identities. [21]
- **Face recognition**: Facial recognition service is a facial recognition service. This suggests that you can't identify unknown faces among the faces collection until you have added recognized faces to it. When inputting an unfamiliar face, the service provides the most similar faces. To confirm the identity of the individual for whom faces have been collected, the face recognition service includes a validation endpoint. [22].

International Scientific Journal of Engineering and Management
Volume: 03 Issue: 05 | May – 2024
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

- **Face verification:** Face verification service verifies if this person is indeed who they say they are. When you send two faces to the rest endpoint, the service compares them and returns how similar they are. [23][24]

4. Result and Discussions

The results of the project is covered in this section. The outcomes of the classification method were reported in this research in terms of F1-Score, recall, accuracy, and precision. Good accuracy is a must for any classification system, and this depends on how effectively picture properties are recovered [25][26].

Accuracy: Accuracy is used to calculate the percentage of data that were successfully categorized. It displays the frequency with which our classifier correctly classifies objects. It is calculated by dividing all values by the sum of all true values.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: The precision of the model indicates its capacity to precisely classify positive instances. This can be computed by dividing the tally of accurately prognosticated affirmative occurrences by the aggregate sum of authentic affirmative occurrences [27].

$$Precision = \frac{TP}{TP + FP}$$

Recall: It is utilized to assess the model's precision in estimating favourable outcomes. Recall is the proportion of veritable positives to the aggregate of true positives and misestimated negatives. [28].

$$Recall = \frac{TP}{TP + FN}$$

F1-Score: It embodies the harmonic mean of precision and recall. It is useful when you need to take both Precision and Recall into account.

International Scientific Journal of Engineering and Management
Volume: 03 Issue: 05 | May – 2024
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Fig. 3 – Sample images

Upon task completion, comprehensive information regarding the confusion matrix is documented, as depicted in DB 1.

Input Images (N) = 10	Sample Person 1	Sample Person 2	Sample Person 3	Sample Person 4	Sample Person 5
Sample Person 1	9	0	0	0	1
Sample Person 2	0	10	0	0	0
Sample Person 3	0	0	10	0	0
Sample Person 4	1	1	0	8	0
Sample Person 5	0	0	0	0	10

DB 1

.

Accuracy, precision, recall, and F1 Score are among the performance measures that are produced based on the confusion matrix computed above.

Performance Metrics (%)	Person 1	Person 2	Person 3	Person 4	Person 5
Accuracy	90.9	100	100	80	100
Precision	90	90.9	100	100	90.9
Recall	90	100	100	80	100
F1-Score	90	95.23	100	88.88	95.23

Table 2 - Performance metrics results in percentage form that are derived from the suggested classification technique.

5. Conclusion

Finally, we have presented a face verification system that employs cutting edge technologies to reliably and quickly use facial recognition to confirm an individual's identification. FaceNet is used for extracting face features, YOLO is used for object recognition, and ArcFace is used for model training. These state-of-the-art technologies allow us to extract high-quality facial features, recognize and extract faces from input images, and train our model for increased resilience and accuracy. The technology, which is applicable to security applications, law enforcement, and access control systems, measures the resemblance of faces using Euclidean distance.

6. References

[1] Insaf Adjabi, Abdeldjalil Ouahabi, Amir Benzaoui, Abdelmalik Taleb-Ahmed, "Past, Present, and Future of Face Recognition: A Review", Journal Electronics 2020, *9*(8), 1188. Published: 23 July 2020.

[2] Yassin Kortli, Maher Jridi, Ayman AlFalou, Mohamed Atri, "Face Recognition Systems: A Survey", Journal, Sensors 2020, 20(2), 342. Published: 7 January 2020.

[3] Murat Taskiran, Nihan Kahraman, Cigdem Eroglu Erdem, "Face recognition: Past, present, and future (a review)", Digital signal processing Journal, Published: 16 July 2020.

[4] Ashu Kumar, Amandeep Kaur and Munish Kumar, "Face detection techniques: a review", Artif Intell Rev 52, 927-928(2019). DOI: 10.1007/s10462-018-9650-2.

[5] Lacopo Masi, Yue Wu, Tal Hassner, Prem Natarajan, "Deep Face Recognition: A Survey", 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images, Parana, Brazil, Doi: 10.1109/SIBGRAPI.2018.00067.

[6] Anagha S. Dhavalikar, R.K. Kulkarni, "Face detection and facial expression recognition system", 2014, International Conference on Electronics and Communication Systems (ICECS), Coimbatore, India, Doi: 10.1109/ECS.2014.6892834.

[7] Madan Lal, Kamlesh Kumar, Rafaqat Hussain, "Study of Face Recognition Techniques: A Survey", June 2018, International Journal Advanced Computer Science and Applications 9(6), DOI: 10.14569/1JACSA.2018.090606.

[8] Muhtahir Oloyede, Gerhard P. Hancke, Herman Myburgh, "Improving Face Recognition Systems Using a New Image Enhancement Technique, Hybrid Features and the Convolutional Neural Network", November 2018, IEEE, DOI: 10.1109/ACCESS.2018.2883748.

[9] Song Zhou, Sheng Xiao, "3D face recognition: a survey", Hum. Cent. Comput. Inf. Sci. 8, 35(2018), Doi: 10.1186/s13673-018-0157-2.

ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

[10] Waqar Ali, Wenhong Tian, Salah Ud Din, Desire Iradukunda, Abdullah Aman Khan, "Classical and Morden Face Recognition Approaches", Multimed Tools Appl 80, 4825-2880 (2021), DOI: 10.1007/s11042-020-0985-1.

[11] Mei Wang, Weihong Deng, "Deep Face Recognition: A Survey", January 2021, DOI: 10/1016/j.neucom.2020.10.081.

[12] Shaveta Dargan, Munish Kumar, Maruthi Rohit Ayyagari, Gulshan Kumar, "A Survey of Deep Learning and Its Applications", Arch computat Methods Eng 27, DOI: 10.1007/s11831019-09344.

[13] Prentza, S. Maglavera, "Delivery of healthcare services over mobile phones". WHO Guidelines for medicine donations

[14] Mohseni, S., Yang, F., Pentyala, S., Du, M., Liu, Y., Lupfer, N., ... & Ragan, E. (2021, May). Machine learning explanations to prevent overtrust in fake news detection. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 15, pp. 421-431).

[15] Pramanik, Sabyasachi, et al. "A novel approach using steganography and cryptography in business intelligence." Integration Challenges for Analytics, Business Intelligence, and Data Mining. IGI Global, 2021. 192-217.

[16] Babu, S. Z., et al. "Abridgement of Business Data Drilling with the Natural Selection and Recasting Breakthrough: Drill Data With GA." Authors Profile Tarun Danti Dey is doing Bachelor in LAW from Chittagong Independent University, Bangladesh. Her research discipline is business intelligence, LAW, and Computational thinking. She has done 3 (2020).

[17] NARAYAN, VIPUL, A. K. Daniel, and Pooja Chaturvedi. "FGWOA: An Efficient Heuristic for Cluster Head Selection in WSN using Fuzzy based Grey Wolf Optimization Algorithm." (2022).

[18] Faiz, Mohammad, et al. "IMPROVED HOMOMORPHIC ENCRYPTION FOR SECURITY IN CLOUD USING PARTICLE SWARM OPTIMIZATION." Journal of Pharmaceutical Negative Results (2022): 47614771.

[19] Mall, Pawan Kumar, et al. "Early Warning Signs Of Parkinson's Disease Prediction Using Machine Learning Technique." Journal of Pharmaceutical Negative Results (2022): 47844792.

[20] Paricherla, Mutyalaiah, et al. "Towards Development of Machine Learning Framework for Enhancing Security in Internet of Things." Security and Communication Networks 2022 (2022).

ISSN: 2583-6129
DOI: 10.55041/ISJEM01802

[21] Tyagi, Lalit Kumar, et al. "Energy Efficient Routing Protocol Using Next Cluster Head Selection Process In Two-Level Hierarchy For Wireless Sensor Network." Journal of Pharmaceutical Negative Results (2023): 665-676.

[22] Sawhney, Rahul, et al. "A comparative assessment of artificial intelligence models used for early prediction and evaluation of chronic kidney disease." Decision Analytics Journal 6 (2023): 100169.

[23] Srivastava, Swapnita, et al. "An Ensemble Learning Approach For Chronic Kidney Disease Classification." Journal of Pharmaceutical Negative Results (2022): 2401-2409.

[24] Mall, Pawan Kumar, et al. "FuzzyNet-Based Modelling Smart Traffic System in Smart Cities Using Deep Learning Models." Handbook of Research on Data-Driven Mathematical Modeling in Smart Cities. IGI Global, 2023. 76-95.

[25] Narayan, Vipul, A. K. Daniel, and Pooja Chaturvedi. "E-FEERP: Enhanced Fuzzy based Energy Efficient Routing Protocol for Wireless Sensor Network." Wireless Personal Communications (2023): 1-28..

[26] Narayan, Vipul, et al. "Enhance-Net: An Approach to Boost the Performance of Deep Learning Model Based on Real-Time Medical Images." Journal of Sensors 2023 (2023)

[27] Narayan, Vipul, et al. "Deep Learning Approaches for Human Gait Recognition: A Review." 2023 International Conference on Artificial Intelligence and Smart Communication (AISC). IEEE, 2023.

[28] Narayan, Vipul, et al. "FuzzyNet: Medical Image Classification based on GLCM Texture Feature." 2023 International Conference on Artificial Intelligence and Smart Communication (AISC). IEEE, 2023