

EVALUATION OF FACE DETECTORS FOR FACE RECOGNITION

M.ZAIN ABBAS¹, JUNAID BABER¹, MAHEEN BAKHTYAR¹, IHSANULLAH¹, WAHEED NOOR¹

¹University of Balochistan, Quetta

*Correspondence: m.zain_malik@yahoo.com

ARTICLE INFORMATION

Crossref Doi:

<https://doi.org/10.56819/pjest.v3i2.77>

Received: 30th-Dec-2022

Revised and Accepted:

13th-January-2023

Published On-Line

21-Jan-2023

ABSTRACT

Artificial intelligence technologies are now advancing swiftly and providing us with a variety of opportunities. The completion of analysis, predicting, and acknowledgment. In recent years, the field of computer vision has become one of highly exciting research. In this paper, we have compared three famous face detectors model, used for face recognition; Dlib, MTCNN, and FaceNet. We have conducted serious of experiments to compare the detection speed and accuracy for above mentioned detectors. FaceNet is the fastest among others without loss in recognition accuracy on commodity hardware.

Keywords: CCTV, DLIB, MTCNN, Face Net, HOG, PCA, CNN, GPU

*Corresponding Author:

Junaid Babar:

m.zain_malik@yahoo.com

Original Research Article



Pakistan Journal Emerging Sciences and Technologies (PJEST) by [Govt. Islamia College Civil Lines Lahore, Pakistan](#) is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).

INTRODUCTION:

Face recognition is the most common problem in face recognition-based applications in computer vision. The most recent technologies created by the major technological corporations in the world also support this. Face ID was presented at an Apple special event in September, active use of facial recognition technology by social media networks to tag users in images.

The human face is the single most crucial factor in determining a person's identity. Biometric traffic strategies over the past two decades have revealed the use of data rules and methods of person identification [1]. One of the most widely used identification systems is thought to be this one. one of the most crucial fields in the identification system It is used to get around issues with traditional identification techniques like: Think about components like your user name, password, and personal identification number (PIN)[2]. As the face is commonly used to evaluate people's personalities in a range of social circumstances, a workable descriptor should be constructed to define which is utilized to execute this system. As a result, the analyst grew interested in facial recognition computations, which are currently among the most well-known topics in the fields of biometrics and computer vision[3]. Face detection in CCTV footage is difficult due to uncontrolled views of things like lighting, look, stance, etc. A shift in lighting can cause a shadow to be thrown across a person's face. The ability to recognize faces accurately can be difficult for a number of reasons, including moving images, jewelry wearers, and cloud

photography[4]. It is possible to determine a person's personality using verifiable evidence. The assertion has been confirmed. A mystery person's personality is investigated, also mentioned here is balanced coordination. Spy cameras, infrared beacons or other techniques

can be used for facial recognition. Face recognition applications are used for character verification, security, monitoring, and crime scene investigation[5]. Even though this subject has seen successful research and numerous computations have been added, additional development may yet be done to boost the productivity of the frameworks because no approach is without flaws[6].

LITERATURE REVIEW

The face recognition stage is used to locate and perceive the human face image in the framework. Any earlier identified inclusion vectors for human faces are eliminated during the element extraction stage. The recognition step finally compares all the layouts of the extracted facial features with the facial data set to discover human face discriminative evidence[7]. The key development in the face recognition framework is the placement of human faces in a given image[8]. This step's goal is to spot any human appearances in the image of the information. Face recognition may be difficult owing to differences in lighting and appearance. The construction of a new face recognition framework is made simpler and more effective by using pre-preparing processes. Viola-Jones-detector [9], histogram of oriented gradient (HOG)[10], and principal component analysis (PCA)[11], are [8][12][11] Only a few strategies were used to find and identify the human facial image. The face detection step can be used, among other things, for object detection, region of interest detection, and categorization of videos and images. [13] Because a wide range of unique arrangements have been offered in the field of acknowledgment, highlight the real techniques and improvements for all phases of the improvement of the acknowledgment framework. [14] Because a wide range of unique arrangements have been offered in the field of acknowledgment, highlight the real techniques and improvements for all phases of the improvement of the acknowledgment framework. [15] A real-time framework for closed-circuit video (CCTV) picture face identification and recognition based on machine learning and deep learning. Traditional CCTV systems are ineffective and costly because a person must be on duty all the time. Since it requires less human input, is less expensive, and can be used to identify suspects, the missing, and anyone accessing a restricted area, the automatic facial recognition system in CCTV footage can be helpful for various businesses, including law enforcement. The problems with image-based recognition include scaling, rotation, crowded backdrops, and variations in light intensity, to name just a few. This paper recommends using a range of feature extraction and face identification methods to construct a CCTV image-based human face recognition system. [15]. The goal of this work is to propose a web-based facial recognition system for use in real-world situations, such online classes. Real-time face recognition is an intriguing field that is getting harder and harder to master because of things like altering illumination, occlusion, changing facial expressions, etc. During the present COVID-19 pandemic scenario, demand for online classrooms has increased significantly. This has increased the need for an efficient, inexpensive, simple, and useful way to track students' attendance in a real classroom. [16] A new examination of the limitations of face detection. They assessed Observer's abilities in four behavioral tests to recognize various faces by differentiating between two different display patterns. In fixed displays, all of the elements were

of the same type (all faces or all non-faces). Mixed displays combined faces and non-faces. Importantly, processing of all items is required for a "fixed" response. We found that adding more faces could be recognized without affecting performance, but that this capacity-free performance was affected by the visual context[16]. This study evaluates the most recent developments in NICU face detection. Since such methods usually fall short in difficult NICU scenarios, we show how fine-tuning can improve neonatal face detector resilience, resulting to our NICU face models. In three studies, the large and diverse neonatal dataset that was used to carry out gold standard face annotations was provided by actual patients admitted to the NICU[17]. The aim of this work is to increase the efficiency of face recognition in dynamic frames. There have been studies that focus on maintaining the high detection rate and enhancing accuracy. The experiment was conducted utilizing the most recent dynamic frame recording of a group of people who were chosen at random to use a treadmill. Here, the subject was viewed from various angles in order to study and extract the variety of face features. The existence of a large number of datasets and cost- effective processing power is subject to value enhancement in the presence of a convolution neural network (CNN) on various object identification and realization criteria. [18]This research presents a real-time framework for cooperative face detection and FLL. Utilizing the overlap between the two goals, we build a fully convolution network to predict the locations of facial landmarks and face regions. Suggested two methods for graph matching without programmable parameters. The entire board maintains the end-to-end property. With regard to accuracy and runtime, extensive testing reveals that our techniques produce state-of-the-art FLL and face detection in the different datasets.

DLIB

The Dlib library is undoubtedly one of the most popular libraries for facial recognition. Face Acknowledgement, a Python program, transforms the Dlib face acknowledgment scheduler into an easy-to-use API. We can more successfully manage issues like acknowledgment as PC vision technology develops quickly. As technology advances, a substantial number of libraries are available to developers to address PC vision issues.

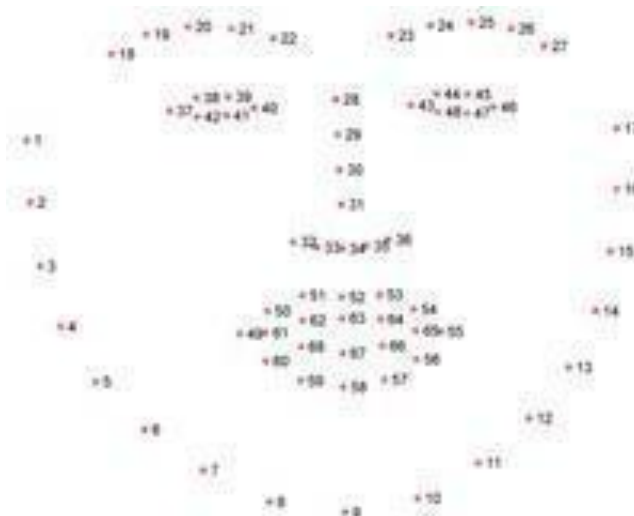


Fig. 1: An Example of Dlib facial landmarks on human face

FACENET

FaceNet uses deep convolution networks rather than go-between bottleneck layers to improve the performance of current deep learning frameworks. This strategy is known as single-shot learning. With this method, the underlying model may be created from a single sample of a facial shot. When new models are made available, the primary model can use them without having to be retrained. Directly in Euclidean space, where facial model similarity is assessed, FaceNet generates a face. Once the benefits of closeness between face models are realized, FaceNet appended with highlight vectors will be essential for face identification proof and grouping. FaceNet uses trios to cooperate with faces during the planning stage by using the online creative trio mining technique. Each of the three anchor photos in this trio is made up of both positive and negative images, indicating that they are part of a series. The three key components of strategies for triplet loss training are anchor, positive, and negative. In this triplet loss, the distance between anchors is positively minimized while the distance between anchors is maximally negative. In contrast to the negative, which has a different identity, this positive has one in common with the anchor [19].

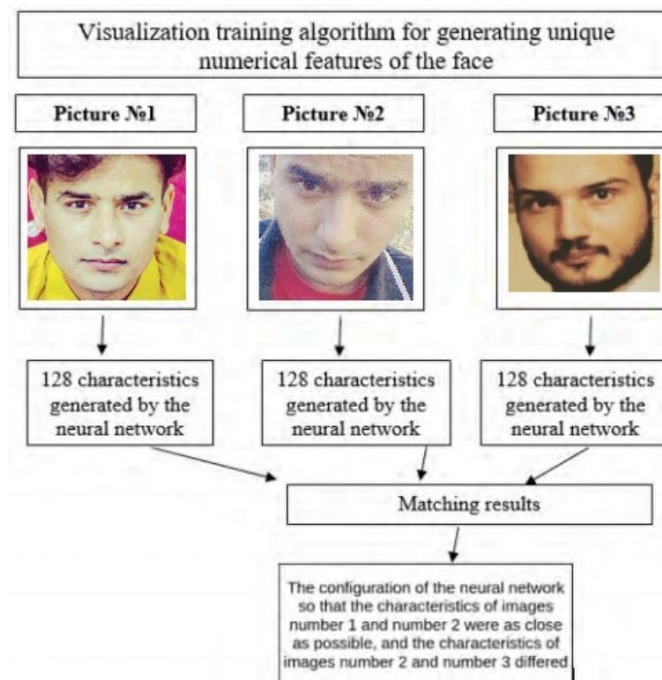


Fig. 2: Abstract flow diagram for Face Net algorithm for generating unique numerical features

MTCCN

An example of a neural network used to assess visual input and images is the convolution neural network (CNN). Perform a variety of tasks MTCNNs, or Multi Task Cascaded Neural Networks, are a more sophisticated form of CNN[19]. The proposed MTCNN model's accuracy is demonstrated in the relationship between disclosure and transformation is peculiar. Utilize outlines to enhance performance. What the project's goals are Avalanches will be created and used as a source of energy in the projected MTCNN. It could be used as a tool to complete a variety of jobs and acquire environmental data. Using a coarse-to-fine process, it can be located on the face. The MTCNN project also aims to combine two tasks[20].

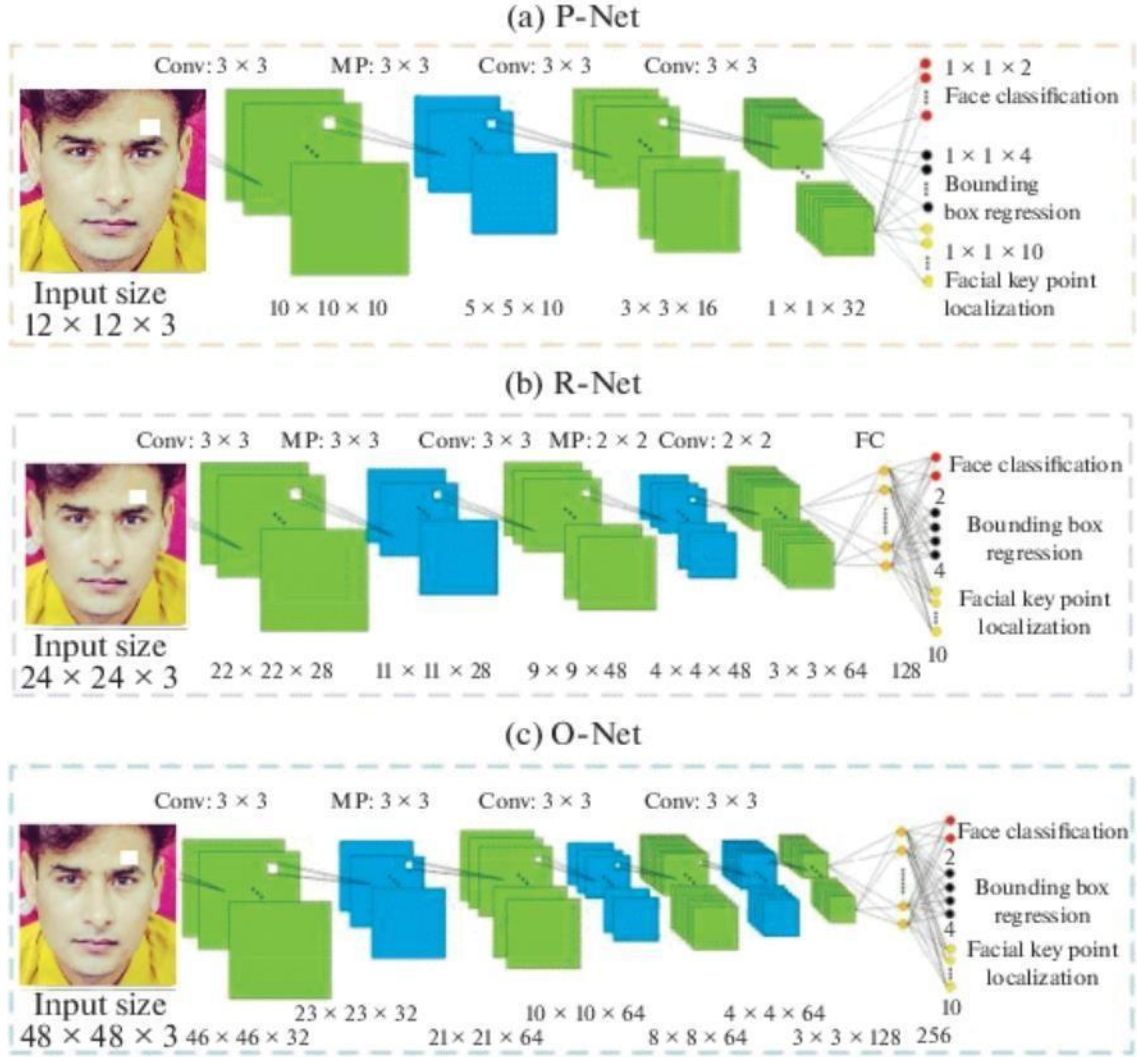


Fig.3: General Architecture of MTCNN

METHODOLOGY

With GPU support empowered, each bundle is evaluated for its speed in perceiving faces in an assortment of 300 photographs. The detection is done at three distinct resolutions. Prior to performance testing, any one-time initialization activities, such as model instantiation, are completed.



Fig. 4: Example of frames captured from video on different time stamps



Fig.5: Faces detected By Dlib



Fig. 6: Faces detected by MTCNN



Fig. 7: Faces detected by FaceNet

RESULT AND DISCUSSION

We tested face detection on three different resolutions to see which one is the fastest one to detect the faces. In our testing we found out that FaceNet was the fastest one in every resolution. It performed very well compare to other face detection models. Below are the Results, as shown in Figure 8.

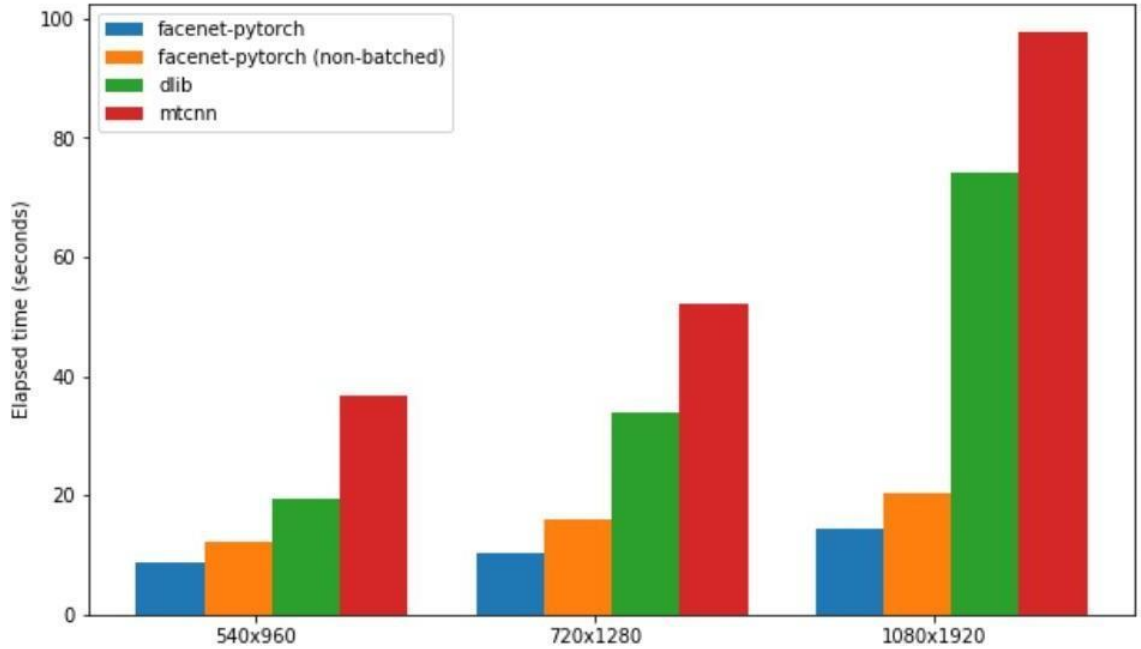


Fig. 8 :Time Comparison of three models Facenet, Dlib, MTCNN.

The tabular representation of experiments are shown below in Table 1.

Table I: Face detection time on different resolution.

Donor Dataset	Purpose		
	(1080x1920)	(720x1280)	(540x960)
MTCNN	13.05	21.04	27.43
DLIB	10.03	15.32	20.13
FaceNet- pytorch(non-batched)	4.02	8.54	14.95
FaceNet- pytorch	3.01	6.20	8.70

Conclusion

Face recognition is widely used for several applications from entertainment to security, and face detection is the first step for face recognition. There are number of detectors available, however, the top 3 are Dlib, FaceNet, and MTCNN which are widely accepted detector for face recognition-based applications. We compared these detectors on famous dataset with different resolution configurations. There are two publicly available implementation of FaceNet; one is with pytorch based. We have used both implementations for our experiments. It can be seen in Table 1, that FaceNet is fastest in all resolutions. Experiments also show that there is no loss in accuracy. The qualitative analysis of face detectors can be seen in Figure 5-7.

Author's Contribution: J.B did the language and grammatical edits, Critical revision manage. M.B, Conceived the idea, designed the work, M.Z.A executed simulated work, data analysis, and interpretation of data. M. Z. A executed and simulated the works, J. B and M. B contributed to supervision and Analysis, I. U and W. N helped in writing and visualization.

Funding: The publication of this article was funded by no one.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgement: The authors would like to thank the department of CS and IT for providing the hardware and assistance with the collection of data.

REFERENCES

- [1] M. Martin, K. Štefan, and F. J. T. Ľubor, "Biometrics authentication of fingerprint with using fingerprint reader and microcontroller Arduino," vol. 16, no. 2, pp. 755-765, 2018.
- [2] M. Pathak, N. Srinivasu, and V. J. T. Bairagi, "Effective segmentation of sclera, iris and pupil in noisy eye images," vol. 17, no. 5, pp. 2346-2354, 2019.
- [3] D. S. Trigueros, L. Meng, and M. J. a. p. a. Hartnett, "Face recognition: From traditional to deep learning methods," 2018.
- [4] D. Kong, Y. Cui, L. Kong, S. J. S. A. P. A. M. Wang, and B. Spectroscopy, "Classification of oil pollutants based on excitation-emission matrix fluorescence spectroscopy and two-dimensional discriminant analysis," vol. 228, p. 117799, 2020.
- [5] R. Jafri and H. R. J. j. o. i. p. s. Arabnia, "A survey of face recognition techniques," vol. 5, no. 2, pp. 41-68, 2009.
- [6] T. Napoléon, A. J. O. Alfalou, and L. i. Engineering, "Pose invariant face recognition: 3D model from single photo," vol. 89, pp. 150-161, 2017.
- [7] A. Vinay, D. Hebbar, V. S. Shekhar, K. B. Murthy, and S. J. P. C. S. Natarajan, "Two novel detector-descriptor based approaches for face recognition using sift and surf," vol. 70, pp. 185-197, 2015.
- [8] H. Yang, X. A. J. J. o. A. Wang, and C. Technology, "Cascade classifier for face detection," vol. 10, no. 3, pp. 187-197, 2016.
- [9] 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, pp. I-I: Ieee.
- [10] J. Rettkowski, A. Boutros, D. J. J. o. P. Göhringer, and D. Computing, "HW/SW Co-Design of the HOG algorithm on a Xilinx Zynq SoC," vol. 109, pp. 50-62, 2017.
- [11] J. H. Shah, M. Sharif, M. Raza, and A. J. I. A. J. I. T. Azeem, "A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques," vol. 10, no. 6, pp. 536-545, 2013.
- [12] H. J. Seo, P. J. I. T. o. I. F. Milanfar, and Security, "Face verification using the lark representation," vol. 6, no. 4, pp. 1275-1286, 2011.
- [13] L. Saranya and K. Umamaheswari, "Multiple Face Analysis and Liveness Detection Using CNN," EasyChair2516-2314, 2021.
- [14] S. J. A. a. S. Reddy Boyapally, "Facial Recognition and Attendance System Using Dlib and Face_Recognition Libraries," 2021.
- [15] R. Ullah *et al.*, "A Real-Time Framework for Human Face Detection and Recognition in CCTV Images," vol. 2022, 2022.
- [16] R. Qarooni, J. Prunty, M. Bindemann, and R. J. C. Jenkins, "Capacity limits in face detection," vol. 228, p. 105227, 2022.

- [17] A. Singh, S. Prakash, A. Kumar, D. J. C. Kumar, C. Practice, and Experience, "A proficient approach for face detection and recognition using machine learning and high-performance computing," vol. 34, no. 3, p. e6582, 2022.
- [18] Z. Pan, Y. Wang, and S. J. S. P. I. C. Zhang, "Joint face detection and Facial Landmark Localization using graph match and pseudo label," vol. 102, p. 116587, 2022.
- [19] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 815-823.
- [20] I. William, E. H. Rachmawanto, H. A. Santoso, and C. A. Sari, "Face recognition using facenet (survey, performance test, and comparison)," in *2019 fourth international conference on informatics and computing (ICIC)*, 2019, pp. 1-6: IEEE.