



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 11    Issue: VIII    Month of publication: Aug 2023**

**DOI: <https://doi.org/10.22214/ijraset.2023.55559>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Face Quality Detection in a Video Frame

T Bhuvaneshwari<sup>1</sup>, N Ramadevi<sup>2</sup>, E Kalpana<sup>3</sup>

<sup>1, 2, 3</sup>Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad-500075

**Abstract:** Face detection technology is often used for surveillance of detecting and tracking of people in real time. The applications using these algorithms deal with low quality video feeds having less Pixels Per Inch (ppi) and/or low frame rate. The algorithms perform well with such video feeds, but their performance deteriorates towards high quality, high data-per-frame videos. This project focuses on developing such an algorithm that gives faster results on high quality videos, at par with the algorithms working on low quality videos. The proposed algorithm uses MTCNN as base algorithm, and speeds it up for high-definition videos. This project also presents a novel solution to the problem of occlusion and detecting faces in videos. This survey provides an overview of the face detection from video literature, which predominantly focuses on visible wavelength face video as input. For the high-quality videos, we will use Face-MTCNN and KLT, for low quality videos we will use MTCNN and KLT. Open issues and challenges are pointed out, i.e., highlighting the importance of comparability for algorithm evaluations and the challenge for future work to create Deep Learning (DL) approaches that are interpretable in addition to Track the faces. The suggested methodology is contrasted with conventional facial feature extraction for every frame and with well-known clustering techniques for a collection of videos.

**Index Terms:** Face quality, video frame, Face detection, Face recognition.

## I. INTRODUCTION

Deep Learning is a subset of Machine Learning (ML) that is based on Artificial Neural Networks (ANNs) with multiple layers, also known as Deep Neural Networks (DNNs). Face detection has been an area of interest to researchers as a way to improve the face recognition accuracy. By filtering out the low-quality images we can reduce various difficulties faced in unconstrained face recognition, such as, failure in face or facial landmark detection or low presence of useful facial information. In last decade or so, researchers have proposed different methods to assess the face image quality, spanning from fusion of quality measures to using learning-based methods. Different approaches have their own strength and weaknesses. But it is hard to perform a comparative assessment of these methods without a database containing wide variety of face quality, a suitable training protocol that can efficiently utilize this large-scale dataset.[1] Deep Learning algorithms can create new features from a small number of characteristics that already exist in the training dataset thanks to feature production automation, which is one of the benefits of Deep Learning. It works well with unstructured data, is affordable, and supports parallel and distributed algorithms: Deep Learning models can be expensive to train, but once they're up and running, they can help businesses cut back on unnecessary spending. Scalability: Due to its capability to quickly and efficiently handle large amounts of data and carry out a variety of computations, Deep Learning is incredibly scalable. A face detection computer programme locates and measures the size of a human face in digital pictures. The foundation of numerous facial analysis techniques, including face alignment, face recognition, face verification, and face parsing, is face detection. Face quality has gained a significant attention in several fields such as health care, personalized marketing, Financial Fraud Detection, Natural Language Processing Fake News Detection. Crowd surveillance is frequently visited public or private spaces, faces can be used to identify and evaluate crowds. Photography is a Face detection current autofocus technique used by some digital cameras. Facial recognition technology is used by mobile apps to identify areas of interest in slideshows. Attendance is to determine who is present, facial recognition is employed. For access management, it is frequently paired with biometric detection. Face recognition is an access control or identity verification solution frequently incorporates a facial recognition technology to identify and confirm a person from a digital image or videoframe [2].

## II. RELATED WORK

A face detection computer programme locates and measures the size of a human face in digital pictures. The foundation of numerous facial analysis techniques, including face alignment, face recognition, face verification, and face parsing, is face detection. Other uses for facial recognition technologies include content-based picture retrieval, video coding, video conferencing, and crowd surveillance. Paul Viola [5] has proposed a system for face detection that can analyse photos very quickly and achieve excellent detection rates.

The first contribution of this paper is a new image representation called an integral image that allows for very fast feature evaluation. The second contribution of this paper is a simple and efficient classifier that is built by selecting a small number of important features from a huge library of potential features using AdaBoost.

#### A. Problem Definition

This paper focuses on extracting significant features of face quality detection in a video frame. In a face quality detection MTCNN algorithm that gives faster results on high quality videos when compare with low quality video and blur video. This analysis on the existence research work will help of face recognition, face detection and face quality in a video frame by using CNN, MTCNN and KLT algorithms.

### III. METHODOLOGY

#### A. Convolutional Neural Networks (CNN)

A convolutional neural network (CNN) is a neural network that has one or more convolutional layers in its structure. Even though segmentation, classification, and image processing are three of the key applications for CNNs, we will concentrate on image processing in this project.

The input for image processing is often an RGB image (a three-color channel matrix made up of Red, Green, and Blue). However, we will just evaluate grayscale images in this instance (two-dimensional matrix, one colour channel).

#### B. Multitask Convolutional Neural Network (MTCNN)

MTCNN stands for Multitask Convolutional Neural Network. It is based on a cascade framework. Below Figure illustrates this method's overall pipeline.

The model first scales down an input image to create an image pyramid, which is the input to the subsequent three-stage cascaded framework.

- 1) *P-Net*: The Proposal Network (P-Net), is a fully convolutional network, aims to produce candidates for face windows and their bounding box regression vectors.
- 2) *R-Net*: The Refine Network (R-Net) receives the output of the P-Net stage as input, performs additional refine and reject more false candidate predictions, and finally performs a non-maximum suppression.
- 3) *O-Net*: The Output Network (O-Net) is similar to the R-Net stage. However, this stage aims to identify face regions with more supervision. Mainly, the network will output the bounding box plus five facial landmarks' locations.[16]

#### C. Kanade-Lucas-Tomasi (KLT)

The Kanade-Lucas-Tomasi (KLT) feature tracker is a method for extracting features. The main objective of the proposal is to address the issue that standard image recognition methods are usually expensive. In order to focus its search on the location that produces the best match, KLT makes use of spatial intensity data. It examines a lot fewer potential matches between the photos than more time-consuming algorithms.[17]

### IV. PROPOSED METHODOLOGY

The algorithm's integrity is maintained by the model, which successfully uses facial detection for high resolution in films while keeping up with real-time, lower-quality movies. Each of the three sub-frameworks in the proposed design recognises and tracks working together to fulfil an objective, faces appear in a stream of photos. (videos). The image is scaled up and down numerous times before being used to distinguish between different face sizes. The P-network (Proposal) then examines the photos to complete the initial detection.

Low detection thresholds cause a large number of false positives even after NMS (Non-Maximum Suppression), even when done on purpose.



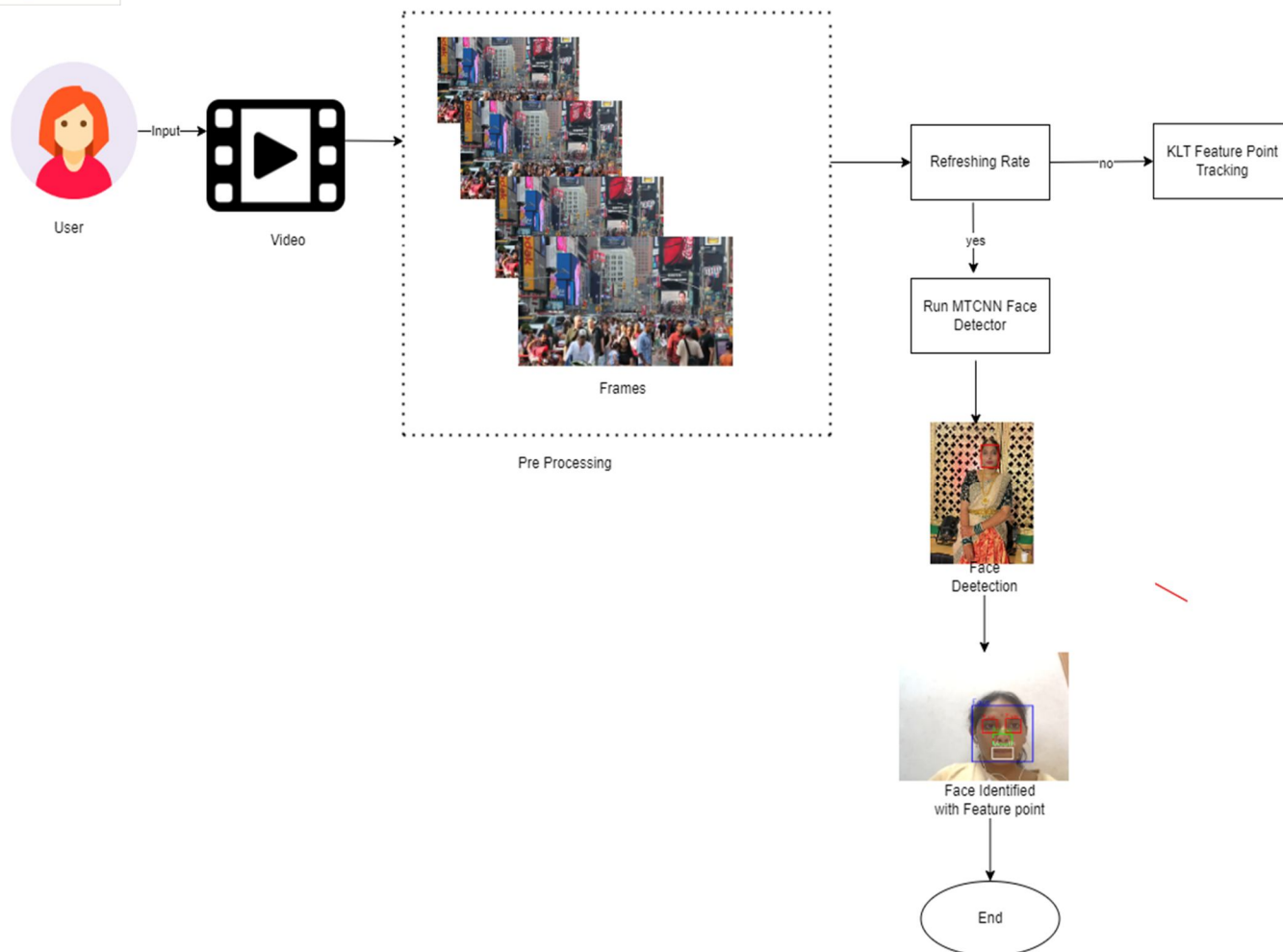


Figure 4.1 Proposed Architecture for face detection

### V. EXPERIMENTS AND RESULTS

In this chapter, we will discuss the outcome of the project.

#### A. High Quality Input Video



Figure 5.1 High Quality Input Video

The above figure represents the high-quality video with 1920\*1080 resolution and 10 seconds video with 50 frame per second. The video is given as input it is converted into ‘n’ number of frames.

*B. Input Video is Converting into Frames*

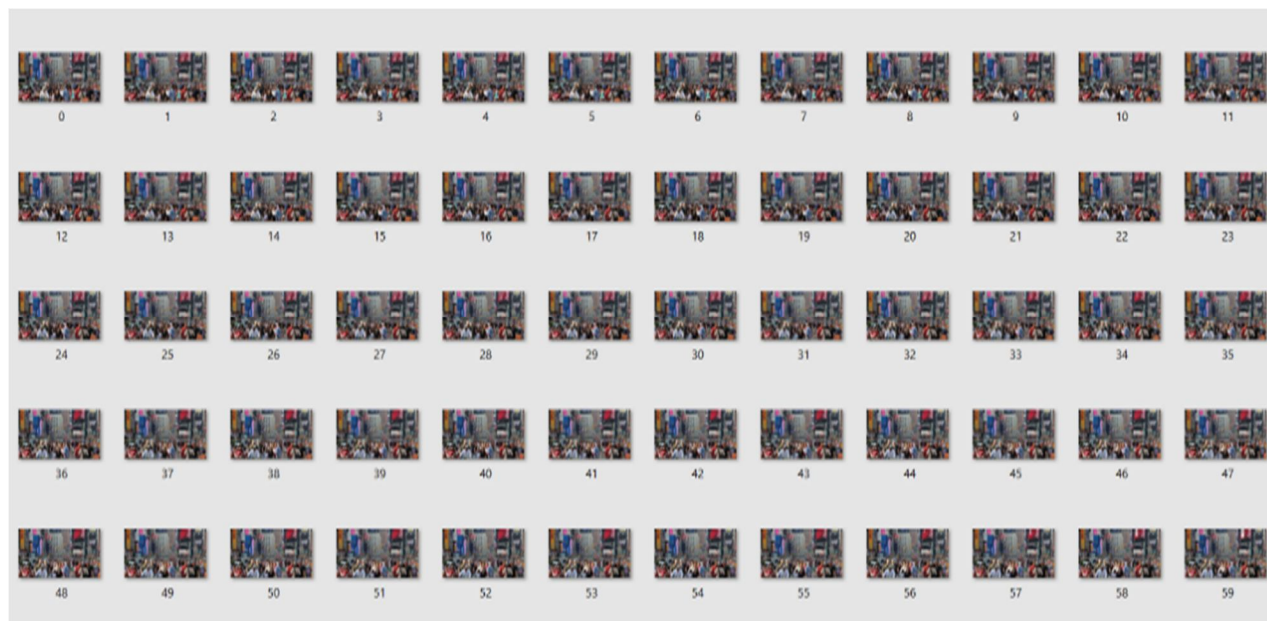


Figure 5.2 High Quality Video Converting into frames

The above figure represents high quality video after we are giving input as video it is going to convert into ‘n’ number of frames. Converting of frames are depend upon the video we are giving.

*C. Face Detection in High Quality Video*

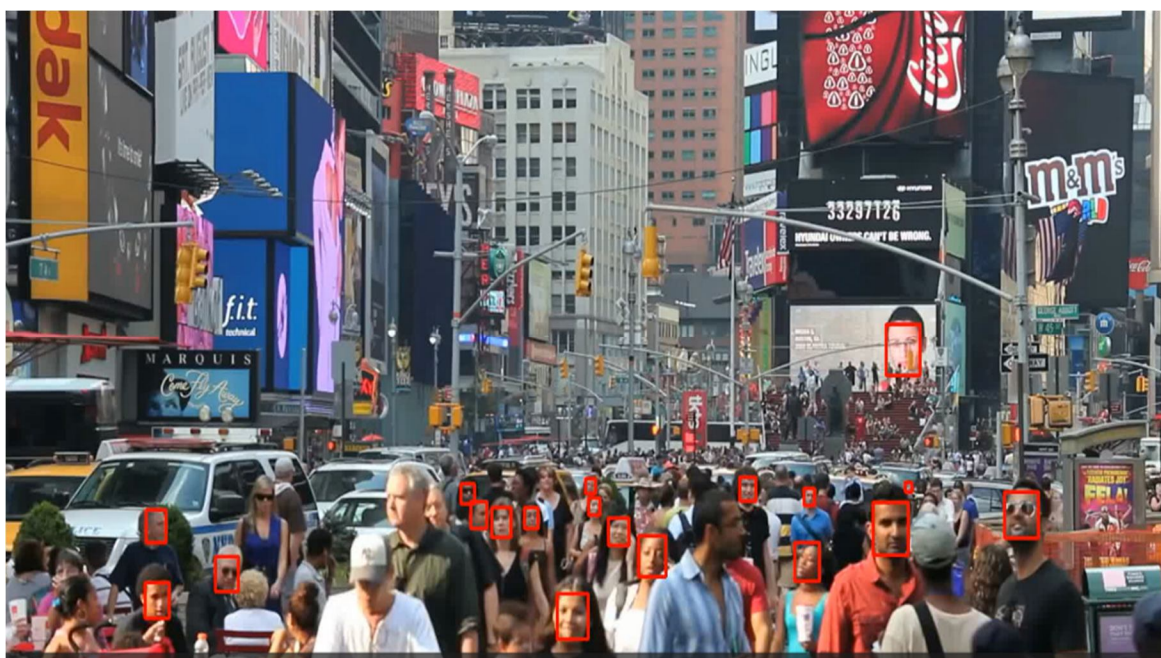


Figure 5.3 Face Detection in high quality video

The above figure shows that detecting the faces of human, after converting the video into frames it is going to detect the human faces and draw a bounding box on the face.



#### D. Detecting Facial Features in High Quality video

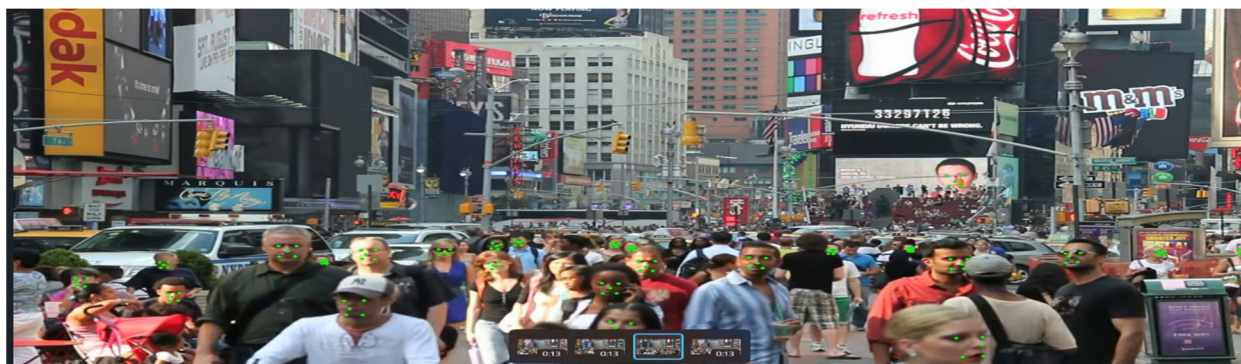


Figure 5.4 Detecting Facial Features in High Quality video

After Detecting the human face by using MTCNN Algorithm it is going to identify the facial feature by using KLT Feature Tracking algorithm.

### VI. CONCLUSION

There are number of reasons that can affect the quality of a face image. These reasons can range from presence of different image sensors, compression algorithms, video or image acquisition conditions, time of acquisition etc. For these varied reasons, automatic face image quality assessment is a very challenging subject. In recent years a number of learning based FIQA methods have been proposed which provides good prediction of face recognition performance based on the face image quality score. This paper deals with three parameters of performance runtime efficiency, accuracy of the detected faces, and occlusion resolution in between frames. Our model achieves optimal performance by trading-off between speed and accuracy. This model can be improved upon by choosing a faster and more accurate base face detection algorithm, since its performance is still dependent on the initial face detecting algorithm chosen for the model.

### VII. FUTURE WORK

In future scope Because of the high performance, face detection method with video segmentation can be widely applied to track faces in massive video data. In order to reduce the false positives in face-like spherical objects and adapt to detect faces in higher resolution video frames, both of video segmentation and face detection procedures need to be further optimized to improve the accuracy of the face detection method.

### REFERENCES

- [1] Wang, Huafeng, Yunhong Wang, and Yuan Cao. "Video-based face recognition: A survey", International Journal of Computer and Information Engineering 3, no. 12 (2009)
- [2] Ismail, Nurulhuda, Mas Idayu Md Sabri. "Review of existing algorithms for face detection and recognition", In 8th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics, pp. 30-39 (2009)
- [3] R. Ranjan et al., "A Fast and Accurate System for Face Detection, Identification, and Verification", in IEEE Transactions on Biometrics, Behaviour, and Identity Science, vol. 1, no. 2, pp. 82-96, April (2019)
- [4] W. Wu, Y. Yin, X. Wang and D. Xu, "Face Detection with Different Scales Based on Faster R-CNN", in IEEE Transactions on Cybernetics, vol. 49, no. 11, pp. 4017-4028, Nov. (2019)
- [5] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." International journal of computer vision 57, no. 2 (2004): 137-154.
- [6] Ismail, Nurulhuda, and Mas Idayu Md Sabri. "Review of existing algorithms for face detection and recognition." In 8th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics, pp. 30-39. 2009.
- [7] Bin Yang, J. Yan, Z. Lei and S. Z. Li, "Aggregate channel features for multi-view face detection," IEEE International Joint Conference on Biometrics, Clearwater, FL, USA, 2014, pp. 1-8, Doi: 10.1109/BTAS.2014.6996284.
- [8] Niu, Gang, and Ququ Chen. "Learning an video frame-based face detection system for security fields." Journal of Visual Communication and Image Representation 55 (2018): 457-463.
- [9] R. Ranjan et al., "A Fast and Accurate System for Face Detection, Identification, and Verification," in IEEE Transactions on Biometrics, Behaviour, and Identity Science, vol. 1, no. 2, pp. 82-96, April 2019, Doi: 10.1109/TBIOM.2019.2908436.
- [10] W. Wu, Y. Yin, X. Wang and D. Xu, "Face Detection with Different Scales Based on Faster R-CNN," in IEEE Transactions on Cybernetics, vol. 49, no. 11, pp. 4017-4028, Nov. 2019, Doi: 10.1109/TCYB.2018.2859482.



- [11] R. Ranjan, V. M. Patel and R. Chellappa, "Hyper Face: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 1, pp. 121-135, 1 Jan. 2019, Doi: 10.1109/TPAMI.2017.2781233.
- [12] R. Abed, S. Bahroun and E. Zagrouba, "Face Retrieval in Videos using Face Quality Assessment and Convolution Neural Networks," 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 2020, pp. 399-405, Doi: 10.1109/ICCP51029.2020.9266253.
- [13] A. Chintha *et al.*, "Recurrent Convolutional Structures for Audio Spoof and Video Deepfake Detection," in IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 5, pp. 1024-1037, Aug. 2020, Doi: 10.1109/JSTSP.2020.2999185.
- [14] Abed, R., Bahroun, S. & Zagrouba, E. "Key Frame extraction based on face quality measurement and convolutional neural network for efficient face recognition in videos". *Multimed Tools Appl* 80, 23157–23179 (2021). <https://doi.org/10.1007/s11042-020-09385-5>
- [15] V. Hosu, H. Lin, T. Sziranyi and D. Saupe, "KonIQ-10k: An Ecologically Valid Database for DL of Blind Image Quality Assessment," in IEEE Transactions on Image Processing, vol. 29, pp. 4041-4056, 2020, Doi: 10.1109/TIP.2020.2967829.
- [16] L. Zhang, H. Wang and Z. Chen, "A Multi-task Cascaded Algorithm with Optimized Convolution Neural Network for Face Detection," 2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS), Shenyang, China, 2021, pp. 242-245, doi: 10.1109/ACCTCS52002.2021.00054.
- [17] T. Xu, D. Ming, L. Xiao and C. Li, "Stitching Algorithm of Sequence Image Based on Modified KLT Tracker," 2012 Fifth International Symposium on Computational Intelligence and Design, Hangzhou, China, 2012, pp. 46-49, doi: 10.1109/ISCID.2012.163.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)