A Comprehensive Survey on Face Quality Detection in a Video Frame

T Bhuvaneshwari[1*], *N* Ramadevi[2], *E* Kalpana[3], *Dr Raman* Dugyala[4] and *G Vijendar* Reddy[5]

[1]Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad-500075

[2]Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad-500075

[3]Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad-500075

[4]Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad-500075

[5]Department of Information and Technology, Gokaraju Rangaraju Institute of Technology, Hyderabad-500090

Abstract. The correctness of the generated face data, which is impacted by a number of variables, significantly affects how well face analysis and recognition systems perform. By automatically analysing the face data quality in terms of its biometric value, it might be able to identify low-quality data and take the necessary action. With a focus on visible wavelength face image input, this study summarises the body of research on the evaluation of face picture quality. The use of DL-based methods is unquestionably expanding, and there are major conceptual differences between them and current approaches, such as the inclusion of quality assessment in face recognition models. In addition to image selection, which is the topic of this article, face picture quality assessment can be used in a wide range of application scenarios. The requirement for comparative algorithm assessments and the difficulty of creating Deep Learning (DL) techniques that are intelligible in addition to providing accurate utility estimates are just a few of the issues and topics that remain unanswered. For each frame, the suggested method is compared to traditional facial feature extraction, and for a collection of video frames, it is compared to well-known clustering algorithms.

Keywords: Face quality, video frame, Face detection, Face recognition

* Corresponding author: talaribhuvana309@gmail.com

1.1 Introduction

Face detection is the first stage of a face recognition system. There has been a lot of research in this field, although the most of it involves images. Human faces can appear in a variety of positions and orientations in video frames, making it difficult for researchers to identify them. In general, there are three main methods for face recognition in videos. It starts with frame-based detection at initially.

Numerous established techniques for still images can be applied throughout this process, including statistical modelling, neural network-based, SVM-based, HMM-based, BOOST-based, and color-based face detection. In addition, integrated detection refers to identifying a face in the first frame and following it throughout the entire sequence. Finally, instead of detecting each frame, temporal approach exploits temporal relationships between the frames to detect multiple human faces in a video sequence. In general, such method consists of two phases, namely detection and prediction and then update-tracking.[1]

Multi-layered artificial neural networks (ANNs) are the foundation of deep learning, a subset of machine learning (ML) (DNNs). To carry out tasks like image recognition, audio recognition, and natural language processing, Deep Learning models must have the capacity to automatically learn features from the input. The three most prevalent DL architectures are recurrent neural networks, feedforward neural networks, and convolutional neural networks. [2].

In a number of fields, such as healthcare, targeted marketing, financial fraud detection, and natural language processing fake news identification, a lot of emphasis has been placed on face quality. Crowds in commonly frequented public and private spaces can be identified and evaluated using faces. Numerous digital cameras employ the well-liked face detection autofocus technique. Key slideshow elements are recognised by mobile apps using facial recognition technology. Attendee tracking and identification are done using facial recognition. For access control, it is frequently used in conjunction with biometric detection. Face recognition is a form of identity verification or access control that frequently employs face recognition technology to identify and authenticate people in digital photographs or video frames. [3]

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. Other uses for facial recognition technologies include content-based picture retrieval, video coding, video conferencing, and crowd surveillance.[4]

1.2 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.

Akshay Mool et al. [5] The proposed approach accelerates for high quality videos and is based on Convolutional-MTCNN. In this study, the problem of occlusion is solved and faces that are either completely or partially obscured in the videos are identified. Angelina Kharchevnikova[6] A lightweight convolutional neural network has been created to evaluate the frame quality using DL methods. We propose extracting knowledge from the clumsy existing Face Q-Net model, for which there is no publicly available training dataset, in order to improve the stage's effectiveness when we assess the frame quality. Rahma Abed [7] to extract keyframes from movies, Convolution Neural Networks with Face Quality Assessment (FQA) have been recommended (CNN). The best face quality frames are then chosen by training a Convolution Neural Network (CNN) in a supervised fashion. Javier Hernandez-Ortega [8] has proposed a Face Q-gen quality assessment method for face recognition based on Generative Adversarial Networks (GANs), is presented in this paper. Gioele Ciaparrone [9] This study demonstrates how proposed the first complete Face-based video retrieval [FBVR] pipeline that can handle big datasets of uncontrolled, multi-shot, multi-person videos. We use FBVR to test the effectiveness of our suggested approach using a dataset that has previously been produced for audio visual recognition. We have a 97.25% mean average precision.

1.3 Problem Definition

The given literature indicates that multiple researchers from various countries have acknowledged conducting considerable work on face quality detection. Modern face quality detection techniques and considerable DL feature extraction are the main topics of this study. The objective of this project is to develop an algorithm that can process high-quality videos as quickly as low-quality live video feeds. The techniques CNN, SVM, Convolutional-MTCNN, and KLT will be used in this overview of previous work to aid in face identification, face quality, and face recognition in a video frame.

2 Modules

2.1 Role of Machine Learning In Face Quality Detection In A Video Frame

The methods used in a face quality detection in a video frame is supervised, unsupervised, classification, regression. Figure 1 represents Taxonomy of Machine Learning algorithms. The use of labelled datasets differentiates supervised learning from machine learning. These datasets are designed to "supervise" algorithms and aid in their accurate prediction of outcomes or data classification. The model can monitor its accuracy and progress over time because the inputs and outputs are labelled. Both classification and regression are supervised learning techniques that can be used with data mining [10]. An algorithm is used to tackle classification issues, such as distinguishing between apples and oranges, and accurately divide test data into multiple categories. Machine learning algorithms can be used to separate spam from your email in the real world and store it in a different folder. Support vector machines, random forests, decision trees, and linear classifiers are a few examples of typical classification methods . Another type of supervised learning called regression employs an algorithm to comprehend the relationship between dependent and independent variables. When predicting numbers based on several data components, such as sales revenue estimates for a certain organisation, regression models are useful. Numerous

regression techniques, such as logistic regression, linear regression, and polynomial regression, are employed. High Error Probability Based on certain outcomes, ML allows us to select an algorithm. To do that, each algorithm must be applied to the results. Period and Space It could take longer than you anticipate to run many ML algorithms. Even the finest algorithm occasionally astounds humans. It will take some time for the system to process large and complicated data sets. Selecting an algorithm Selecting an algorithm is still a manual procedure in machine learning. All algorithms need to be evaluated on our data. After that, we can select the algorithm of our choice. Depending on how precisely the outcomes come out, we choose them.[11]
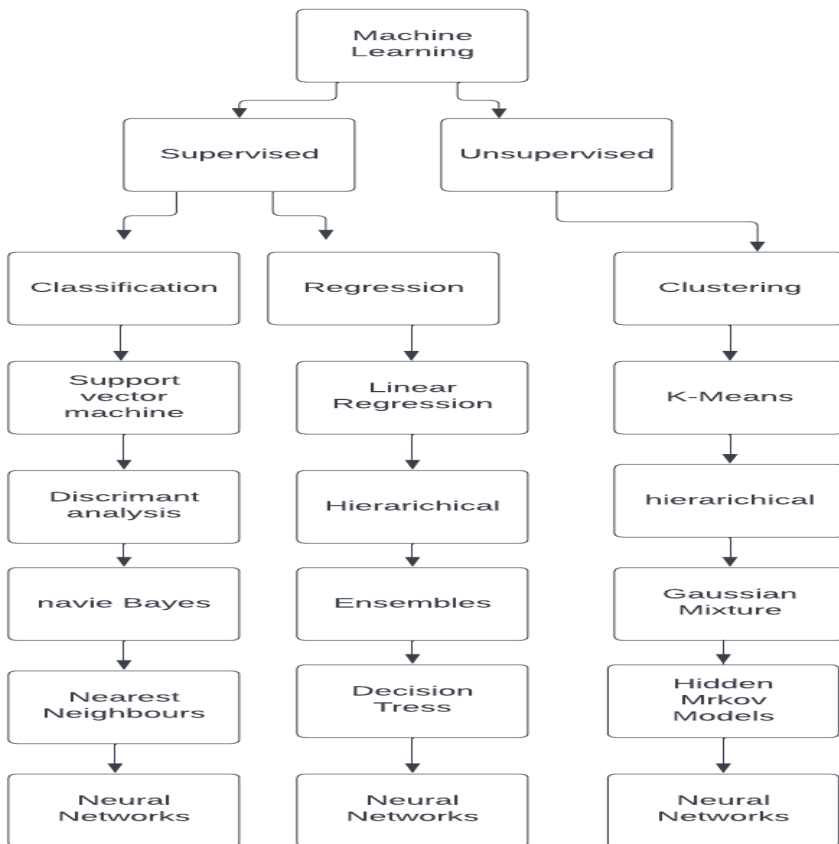

Fig.1. Taxonomy of Machine Learning algorithms

2.2 Role of Deep Learning In A Face Quality Detection In A Video Frame

The methods used in a face quality detection in a video frame is Deep Neural Network, convolutional Deep Neural Network, Recurrent Neural Network. The human brain served as an inspiration for the development of deep neural networks. The Deep Neural Network software predicts and offers solutions by going far beyond the "if and else" conditions. Programming and coding are not necessary with Deep Neural Network AI to obtain the output [12]. Maximising the use of unstructured data Deep learning algorithms can be trained

on various types of data while still producing results that are useful for achieving the training's objectives. Deep learning algorithms, for instance, can be applied to identify any relationships between industry assessments, social media activity, and more to project future stock prices of a specific company. Feature engineering is no longer necessary Since it improves accuracy and occasionally necessitates specialised problem-domain knowledge, feature engineering is a crucial machine learning task. The deep learning approach's ability to execute feature engineering independently is one of its main advantages. decrease in unnecessary spending Because they are so expensive, recalls can cost a company millions of dollars in some industries. Deep learning's drawbacks include For a deep learning system to produce the greatest results, it must be trained on enormous amounts of data. computational capability is the increase in computer capacity that allows us to handle more data is another crucial factor in the popularity of deep learning. The popularity of deep learning has increased due to algorithms. The main reason for these most recent algorithmic developments is that algorithms now run much quicker than they did in the past, allowing for the use of an expanding amount of data. Marketing: The marketing industry has been crucial. The popularity of neural networks has fluctuated over the years, reaching peaks and dips.Fig 2 represents technquies used in deep learning.[13]
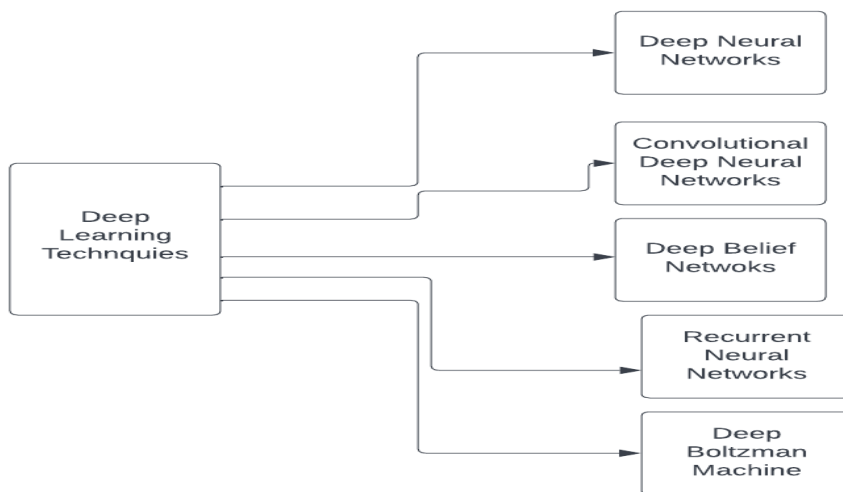
Fig. 2. Techniques of Deep Learning

3 Proposed Methodology

The model utilizes facial detection successfully for high resolution films while maintaining the algorithm's integrity and keeping up with live, lower quality films. The suggested architecture consists of three sub-frameworks, each of which identifies and tracks faces in a stream of images while working in tandem to complete a specific task. (videos). To distinguish between various face sizes, the image is first scaled up and down many times. To complete the first detection, the P-network (Proposal) then looks over the images. Low detection thresholds result in numerous false positives even after NMS (Non-Maximum Suppression), even if this is done on purpose.[1] Fig 3 Represents the flow chart of proposed Method. Fig 3 represents flow chart of proposed method.

1.To separate the frame work we will use Convolutional neural networks [CNN] algorithm.

2. Kanade-Lucas-Tomasi [KLT] Algorithm is used to extract the features.

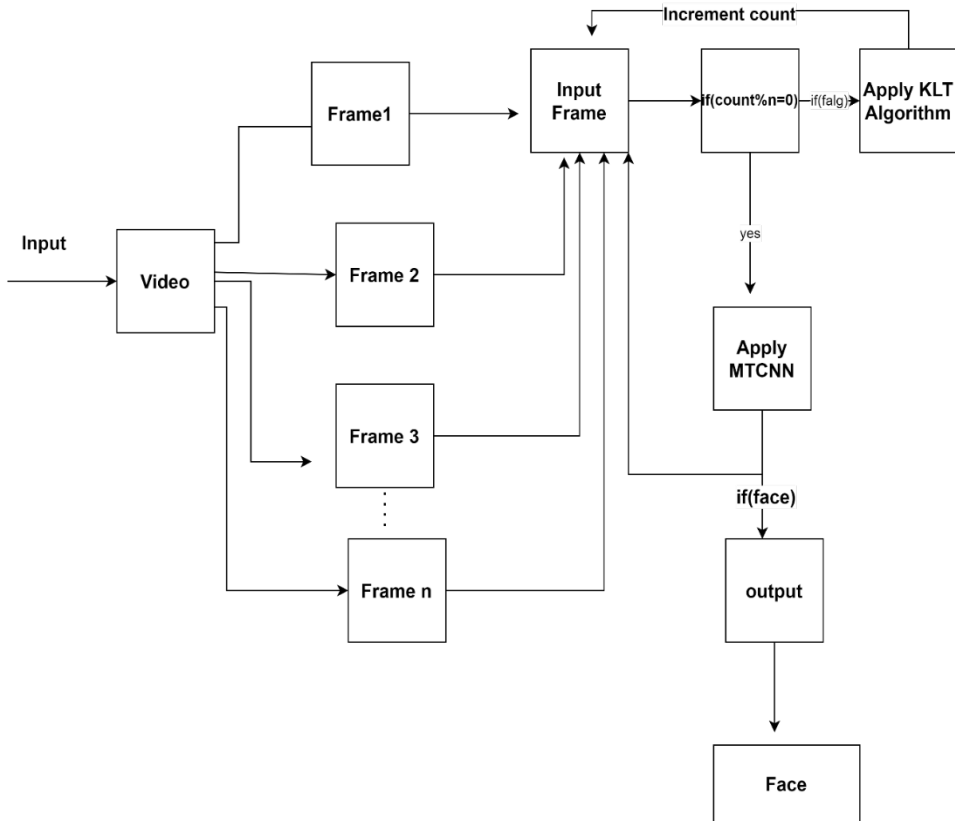3.Multitask Cascaded Convolutional Networks [MTCNN] is used to detect the face.

Fig 3: Flow chart of proposed Method

4 Conclusion

The Several variables can affect how well a face image is captured. These justifications may be brought on by the employment of several picture sensors, different compression techniques, undesirable video or image capture settings, timing difficulties, etc. For all of these reasons, determining the quality of a facial photograph automatically is an extremely difficult problem. Several learning-based FIQA approaches have recently been reported that use the facial image quality score to properly forecast how well face recognition would work. Runtime effectiveness, face detection accuracy, and occlusion resolution between frames are the three performance measures that are discussed in this study. Our model performs best when speed and precision are balanced. This can be done by employing an improved version

of the basic face detection model, which is speedier and more accurate. Consequently, the model's success still depends on the initial face detection technique picked for it.

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. Mool, Akshay, J. Panda, and Kapil Sharma. *Optimizable face detection and tracking model with occlusion resolution for high quality videos*. Multimedia Tools and Applications 81, no. 8 (2022)

6. Kharchevnikova, Angelina, and Andrey V. Savchenko. *Efficient video face recognition based on frame selection and quality assessment*, PeerJ Computer Science 7 (2021)

7. 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)

8. J. Hernandez-Ortega, J. Fierrez, I. Serna and A. Morales*, FaceQgen: Semi-Supervised DL for Face Image Quality Assessment*, 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021), Jodhpur, India, (2021)

9. Ciaparrone, Gioele, Leonardo Chiariglione, and Roberto Tagliaferri. *A comparison of DL models for end-to-end face-based video retrieval in unconstrained videos*. Neural Computing and Applications 34, no. 10 (2022)

10. Kharchevnikova, Angelina, and Andrey V. Savchenko. *Efficient video face recognition based on frame selection and quality assessment*. PeerJ Computer Science 7 (2021)

11. Ou, Fu-Zhao, Xingyu Chen, Ruixin Zhang, Yuge Huang, Shaoxin Li, Jilin Li, Yong Li, Liujuan Cao, and Yuan-Gen Wang. "SDD-FIQA: *unsupervised face image quality assessment with similarity distribution distance*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7670-7679. (2021)

12. Huang, Yu-Hsuan, and Homer H. Chen. *Deep face recognition for dim images. Pattern Recognition* 126 (2022)

13. Ciaparrone, Gioele, Leonardo Chiariglione, and Roberto Tagliaferri. *A comparison of DL models for end-to-end face-based video retrieval in unconstrained videos*. Neural Computing and Applications 34, no. 10 (2022)