A Review on Comparative Analysis of Face Detection Algorithms

Vandna Kumari
Student Department of CSE
IET Bhaddal Technical Campus Ropar
Punjab, India

Barinderjit Kaur
HOD Department of CSE
IET Bhaddal Technical Campus Ropar
Punjab, India

ABSTRACT

Facial recognition plays a vital role in computer vision applications. Several face detection algorithms have been developed over time to accurately detect human faces in images and videos. In this review paper, we present an overview and comparative analysis of traditional face detection algorithms such as Haar cascades and Viola-Jones, as well as newer methods such as SIFT, SURF, ORB, and LBP. We discuss the key features, benefits, and limitations of each algorithm and provide a detailed comparison table for ease of reference. Our analysis shows that each algorithm has strengths and weaknesses, and the choice of algorithm is dependent on the application's specific requirements. We conclude by highlighting the need for more robust and efficient algorithms. Overall, this review paper provides a comprehensive guide for face detection researchers and practitioners.

General Terms

Face Detection, Image processing, Performance, Accuracy, Precision, Recall, F1-Score.

Keywords

Computer vision, Haar cascade classifier, Viola-Jones, SIFT, SURF, ORB, Local Binary Pattern cascade classifier.

1. INTRODUCTION

The technique of discovering and recognizing human faces in photos or videos is known as face detection [1]. Face detection is a fundamental topic in computer vision that has gotten a lot of interest because of its wide variety of applications, which include face identification, biometrics, surveillance, and human computer interaction [2]. As face identification algorithms improve, there is a rising need to evaluate their performance, accuracy, and applicability in various contexts. Comparing face identification algorithms using photos can give useful information about their strengths, limits, and relative performance. This Review Paper provides a comprehensive comparative analysis of image-based face identification methods. The main objective is to compare and evaluate the performance of different algorithms in terms of accuracy, speed, robustness, and adaptability to changing environmental conditions. The aim is to provide academics and practitioners with a clear understanding of the strengths and limitations of these algorithms, enabling informed decision-making for specific applications.

The results of this review paper will add to the existing body of knowledge by offering insights into the comparative performance of image-based face detection systems. The complete evaluation will aid researchers and practitioners in computer vision and related disciplines by allowing them to make proficient decisions about the selection and use of face detection algorithms for specific applications.

The remaining parts of the paper are organized as follows: Section 2 provides a literature review. Section 3 provides an overview of face detection algorithms. Section 4 provides a comparison of these face detection algorithms. Section 5 presents the paper's conclusion.

2. LITERATURE REVIEW

Several studies have focused on the comparative analysis of face detection algorithms, including Viola-Jones, Haar Cascade, LBP, SIFT, SURF and ORB, using precision, recall, F1-score, accuracy, and execution time metrics. These comparative analyses have contributed to understanding the strengths and weaknesses of these algorithms and guiding researchers and practitioners in selecting appropriate algorithms for face detection tasks.

Viola and Jones (2001) developed the Viola-Jones algorithm, which uses a boosted cascade of simple features to recognize objects quickly. Because of its excellent accuracy and real-time processing capabilities, this method has been frequently used for face detection. The study used precision and recall measurements to demonstrate the algorithm's performance [1]. Lienhart and Maydt (2002) extended the Viola-Jones algorithm with the Haar Cascade classifier, which improves both accuracy and efficiency. Their comparative analysis showed that the Haar Cascade classifier achieved higher accuracy and faster execution times compared to the original Viola-Jones algorithm [3]. In a study by Ojala et al. (1996), a comparative analysis of texture measures, including LBP, was conducted for face detection. The study focused on the robustness of texture-based methods to variations in lighting conditions. The results demonstrated the effectiveness of LBP in handling illumination changes and its potential for face detection tasks [4].

The performance of feature-based algorithms such as SIFT and ORB has also been evaluated in comparative studies. Lowe (2004) presented a distinctive image feature analysis, including SIFT, and demonstrated its robustness to scale, rotation, and affine transformations. Rublee et al. (2011) proposed ORB as an efficient alternative to SIFT or SURF, showcasing its computational efficiency and accuracy for feature detection [5, 6, 7].

To further evaluate these algorithms, performance metrics such as precision, recall, F1-score, accuracy, and execution time have been utilized. Ristani et al. (2016) and Kang and Lee (2019) have provided comprehensive surveys on performance evaluation measures for object detection and tracking algorithms, which can be adapted for evaluating face detection algorithms as well [8, 9]. Zhang, Wang, and Li (2019) compared face detection algorithms such as Viola-Jones, Haar Cascade Classifier, LBP, SIFT, and ORB with a focus on video surveillance systems. The algorithms were tested using the Caltech Faces dataset and the FDDB dataset. According to this study, Viola-Jones had the best overall performance, with high precision, recall, F1 score, and accuracy. The Haar Cascade Classifier performed admirably as well. LBP, SIFT, and ORB have lower accuracy but faster execution times [10].

In this study [11], the performance of five face identification algorithms is compared: Viola-Jones, Haar cascade classifier, LBP, SIFT, and ORB. The researchers evaluated the algorithms' performance using the Labeled Faces in the Wild (LFW) dataset. The most accurate algorithm was discovered to be the Viola-Jones algorithm, followed by the Haar cascade classifier, LBP, SIFT, and ORB. Dong, Y., Zhang, Y., Jiang, H., and Chen, Y. (2020) compared the Viola-Jones, Haar Cascade, LBP, and SIFT algorithms for face detection. The study discovered that the Viola-Jones algorithm achieved high precision and recall rates, showing its accuracy in detecting faces. The Haar Cascade algorithm performed well in terms of precision and recall, but the LBP and SIFTS algorithms performed poorly. In terms of execution time, the SIFT method performed the slowest, whereas the Viola-Jones and Haar Cascade algorithms performed more quickly [12].

This study analyses the performance of three different image processing-based face recognition algorithms: Viola-Jones, Haar cascade classifier, and LBP. The LFW dataset was used in the study to assess the performance of the algorithms. The Viola-Jones method had the highest accuracy, followed by the Haar cascade classifier and LBP, according to the study's findings [13].

While previous research has conducted comparative evaluations of these face identification algorithms, it is important to remember that the effectiveness and efficiency of these algorithms can vary depending on the unique dataset, application needs, and implementation specifics. As a result, the goal of this research is to add to the existing body of knowledge by conducting a comprehensive comparison analysis using precision, recall, F1-score, accuracy, and execution time metrics, with a specific focus on face detection tasks. By providing a comprehensive evaluation of these algorithms on images, this study aims to provide substantial insights into their relative performance and assist researchers and practitioners in selecting the best algorithm for their specific needs.

3. FACE DETECTION TECHNIQUES

There are several face detection techniques [14] available. The following face-detection techniques are discussed in this study:

3.1 Face Detection Using Viola – Jones Algorithm

Paul Viola and Michael Jones proposed the Viola-Jones algorithm in 2001. It is a machine-learning based technique for object recognition that makes use of Haar-like features. Before training a classifier with AdaBoost, the method extracts Haar-like features from a picture. The AdaBoost method combines weak classifiers to produce a powerful classifier capable of recognizing objects in photographs [1].

The Viola-Jones algorithm consists of three main stages:

3.1.1 Haar-like Feature Selection

In this stage, a set of Haar-like characteristics is chosen to represent various facial features such as edges, lines, and corners. These are simple rectangular filters that are applied to the image at various scales and places [1].

3.1.2 AdaBoost Learning

In this stage, a machine learning technique called AdaBoost is used to select the most discriminative features from the set of Haar-like characteristics. AdaBoost assigns weights to features based on their ability to correctly categorize faces. The method iteratively selects a collection of features that provide the greatest overall face detection performance [1].

3.1.3 Cascading Classifier

The cascade classifier is a technique for quickly evaluating Haar-like features and rejecting non-face regions. It is divided into several stages, each containing a collection of weak classifiers. The cascade structure enables early rejection of non-face regions, minimizing the computation required for the following phases [1].

It is time-consuming and wasteful to apply 6,000 features to one image; thus, the researchers devised the concept of a cascade classifier. If the window fails in the first stage, subsequent stages become unnecessary and are discarded. If a window passes the first stage, the algorithm moves to the second stage. If the window successfully passes all stages, it is labeled as a face region. This is how the Viola-Jones facial detection algorithm operates [15].

The combination of these stages allows the Viola-Jones algorithm to achieve fast and accurate face detection. It has been widely used in various applications and has set the foundation for many subsequent face detection algorithms [16].

3.2 Face Detection Using Haar Cascade Classifier

The Haar Cascade classifier is a machine learning-based approach for recognizing objects, notably faces.

It detects the existence of an object in a picture by employing a set of Haar-like features. These characteristics are simply rectangular filters that detect changes in the intensity of neighboring pixels. The Haar Cascade classifier is trained on a large set of positive and negative images to grasp the structures that represent the object of interest. The classifier works by swiping the Haar-like filters across the picture at various scales and locations and then deciding whether every region of the picture is similar to the learned patterns. The classifier generates a set of rectangles representing the locations of the discovered objects [16].

The Viola-Jones face detection algorithm recommends the Haar cascade classifier. This technique requires a large number of photos, both positive and negative, to train the classification algorithm. Positive images have faces, whereas negative images do not [17]. Haar Cascade classifiers are based on Haar-like characteristics but recognize objects using a cascade architecture rather than Adaboost. In this method, the image is processed through a succession of phases, with each stage comprising a classifier that gradually eliminates non-object portions of the image [1, 3].

Haar Cascade classifiers are widely used in a variety of applications, including face detection, object detection, and pedestrian detection. There are some features in the Haar cascade classifier: the edge feature, the line feature, and the four-rectangle feature.

There are several types of Haar cascade classifiers used for face detection, including:

Frontal face classifier: This classifier has been trained to detect frontal faces in images.

Profile face classifier: This classifier has been trained to detect profile faces in images.

Eye classifier: This classifier has been trained to detect eyes in images.

Nose classifier: This classifier has been trained to detect noses in images.

Mouth classifier: This classifier has been trained to detect mouths in images.

Full body classifier: This classifier has been trained to detect full bodies in images.

If the window fails in the first stage, subsequent stages become unnecessary and are discarded. If a window passes the first stage, the algorithm moves to the second stage. If the window successfully passes all stages, it is labeled as a face region. This is how the Viola-Jones facial detection algorithm operates [15].

3.3 Face Detection Using Local Binary Pattern Cascade Classifier

Local Binary Pattern (LBP) is a texture descriptor used in image analysis and computer vision applications to identify faces using the LBP feature extraction technique. Ojala et al. presented it in 1996 as a simple and efficient method for texture classification. LBP encodes an image's local structure by comparing the central pixel value to the pixels surrounding it. The LBP operator converts the image into a binary pattern picture, which may be used to encode texture local patterns [18, 4]. LBP divides the image into smaller sub-regions and applies a series of classifiers to each. Each sub-regions LBP feature is extracted, and a binary evaluation is performed to determine whether or not the sub-region contains a face. LBP cascade classifiers work in steps, each with a collection of weak classifiers.

LBP consists of the following stages:

3.3.1 Image Preprocessing

The input image is preprocessed to enhance the quality and remove noise. This may include gray scale conversion, histogram equalization, or image resizing [18, 4].

3.3.2 LBP Feature Extraction

The LBP operator is applied to the preprocessed image to extract local texture features. For each pixel, the binary code is computed by comparing its intensity value with the surrounding neighbors. These codes are then concatenated to form a feature vector that represents the texture information of the image [18, 4].

3.3.3 Training a Classifier

A machine learning classifier such as Support Vector Machines (SVM) or AdaBoost is trained using labeled face and non-face samples. The extracted LBP feature vectors are input to the classifier, which learns to distinguish non-face patterns [19].

3.3.4 Face Detection

Once the classifier is trained, it can be applied to new unseen images for face detection. The LBP features are extracted from the photo, and the classifier predicts whether each region corresponds to a face or non-face. This process is usually performed by sliding a window across the image at multiple scales to detect faces of different sizes [19].

3.3.5 Post-processing

Detected face regions may undergo post-processing steps such as non-maximum suppression to remove overlapping detections or additional filtering to improve the accuracy of the results [19].

These classifiers are trained to detect specific visual features, such as edges and corners, which are combined to create a robust classifier. Each phase in the algorithm produces an output, which is used to decide whether the image should proceed to the next stage or be rejected as a non-face [19]. This implies that as the number of positive images increases, the e-strategy improves.

3.4 Face Detection Using SIFT (Scale-Invariant Feature Transform)

SIFT is a method for recognizing faces in photographs by extracting and matching distinguishing features. The SIFT method discovers and describes local features in images, which are then matched to determine whether or not they are faces. SIFT-based face detection is resistant to changes in position, scale, and lighting. It has been shown to recognize faces with high accuracy in a number of pictures, such as those with complex backgrounds [5].

David Lowe created the SIFT technique in 1999, and it has since been widely applied in a number of computer vision applications such as object detection, picture retrieval, and panorama stitching. Face recognition derives SIFT features from an image and selects a collection of features that are likely to belong to a face. These traits are then matched to a set of predefined face templates to assess whether a face is present in the image [6].

The SIFT algorithm is widely used in computer vision applications such as image registration, object recognition, and three-dimensional reconstruction. However, because of its computational complexity, it is too slow for real-time applications. As a result, re-searchers have created SIFT algorithm variants such as SURF and ORB that have faster computation times while maintaining good performance [7].

3.5 Face Detection Using SURF (Speeded Up Robust Features)

The SURF algorithm (Speeded Up Robust Features) is a popular feature extraction algorithm in computer vision and image processing. It is a SIFT extension that aims to be faster and more robust than SIFT. The SURF algorithm, like SIFT, extracts key points from images and creates descriptors for each key point based on the local picture gradient. SURF, on the other hand, employs a faster method for identifying key points and computing identifiers, making it more suitable for real-time applications [6, 7].

SURF has been used in a variety of computer vision applications, including object recognition, image stitching, and facial recognition. Its robustness and efficiency make it a popular choice for applications that require feature matching in real-time or near real-time. However, like SIFT, SURF is a patented algorithm, and its implementation in commercial applications could require a license from the patent holder.

3.6 Face Detection Using ORB (Oriented FAST and Rotated BRIEF)

ORB (Oriented Fast and Rotated Brief) is a feature detection and presentation approach for face detection. Rublee et al. (2011) proposed ORB as an efficient alternative to SIFT or SURF, showcasing its computational efficiency and accuracy for feature detection [6, 7]. ORB is a SIFT algorithm variant that provides fast computation time while maintaining good performance. ORB discovers and describes picture characteristics by combining the FAST (Agile Segmentation Test Characteristics) algorithm and the BRIEF (Binary Robust Independent Elementary Features) descriptor. The FAST method finds critical points, and the BRIEF descriptor characterizes them. To ensure rotation-invariant descriptions, ORB also contains an orientation assignment step [7]. ORB has been proven to deliver accurate and rapid results for face detection.

It is used in various applications, including security and surveillance systems and mobile devices. Face detection using ORB requires identifying key regions in an input image and computing their descriptors. These key regions and descriptors are then matched to a database of face-related key regions and descriptors. If a match is found, the region surrounding the matched key points is regarded as a face [7].

4. COMPARISON OF FACE DETECTION ALGORITHMS

Table 1. Summarizes the pros and cons of face detection algorithms

Algorithm	Pros	Cons
Haar Cascade	Fast and efficient	Requires a large amount of training data
Viola-Jones	Fast and accurate	Sensitivity to lighting conditions and background clutter
SIFT	Robust to scale And rotation changes	Computationally expensive
SURF	Robust to scale Rotation changes	Computationally expensive
ORB	Fast and efficient	Less robust to illumination changes compared to SIFT and SURF
LBP	Fast and efficient	Less robust to scale and rotation changes compared to SIFT and SURF

Table 1 provides the strengths and weaknesses of each algorithm, enabling users to choose the most suitable approach based on their specific requirements and constraints.

Table 2. Comparison of face detection algorithms based on key parameters

Algorithm	Detection Time	Accuracy	Robustness
Haar Cascade	Fast	Moderate	Sensitive to lighting and pose changes
Viola-Jones	Fast	High	Sensitive to occlusion
SIFT	Fast	High	Robust to changes in scale and orientation
SURF	Fast	High	Robust to changes in scale and orientation
ORB	Fast	High	Robust to changes in scale and orientation
LBP	Fast	Moderate	Robust to changes in lighting and facial expression.

Table 2 provides a comparative analysis of the face detection algorithms based on the specified approach, showcasing their detection time, accuracy, performance, robustness.

5. CONCLUSION

In the digital world of today, authentication and identification have become key challenges. Face detection is important in authentication and identification. There are various established techniques for doing so. After analyzing the numerous face detection techniques, such as Haar Cascades, Viola-Jones, SIFT, SURF, LBP, and ORB, we can conclude that each approach has its own set of benefits and drawbacks. Haar Cascades and Viola-Jones are fast and efficient, but they struggle to distinguish faces in low-light circumstances. SIFT and SURF recognizes faces well under varying situations but are computationally expensive. LBP and ORB are speedier alternatives to SIFT and SURF, however, they may not detect faces as effectively in tough situations.

To summaries, no single algorithm can work flawlessly in all secnaris. The algorithm chosen will be determined by the Application and the trade-offs between speed, accuracy, and computer resources. Overall, face detection is important in a variety of applications, such as security, surveillance, and facial identification. The capacity to detect faces properly and effectively is critical in these applications, and the development of improved face identification algorithms has evolved significantly in recent years.

Based on the analysis, researchers and practitioners can make informed decisions when selecting a face detection algorithm for their specific applications. They can choose the algorithm that best meets their demands by taking into account their individual requirements, priorities, and limits.

6. ACKNOWLEDGMENTS

Sincere gratitude to the professionals and mentors who have provided direct and indirect assistance in the creation of this paper. The insightful feedback, constructive recommendations, and constant encouragement received throughout this endeavor are deeply appreciated.

7. REFERENCES

[1] Paul Viola and Michael 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, volume 1, pages I–I. Ieee, 2001.

[2] Arun A Ross, Karthik Nandakumar, and Anil K Jain. Hand- book of multibiometrics, volume 6. Springer Science & Busi- ness Media, 2006.

[3] Rainer Lienhart and Jochen Maydt. An extended set of haar- like features for rapid object detection. In Proceedings. In- ternational conference on image processing, volume 1, pages I–I. IEEE, 2002.

[4] Ojala, T., Pietikäinen, M. and Harwood, D., 1996. A comparative study of texture measures with classification based on featured distributions. Pattern recognition, 29(1), pp.51-59.

[5] Cong Geng and Xudong Jiang. Face recognition using sift features. In 2009 16th IEEE international conference on im- age processing (ICIP), pages 3313–3316. IEEE, 2009.

[6] David G Lowe. Distinctive image features from scale-invariant key points. International journal of computer vision, 60:91–110, 2004.

[7] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In

2011 International conference on computer vision, pages 2564– 2571. Ieee, 2011.

[8] Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara, and Carlo Tomasi. Performance measures and a data set for multi-target, multi-camera tracking. In Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands,October 8-10 and 15-16, 2016, Proceedings, Part II, pages 17–35.Springer, 2016.

[9] Kang, K., & Lee, Y. (2019). Object Detection and Tracking Performance Measures: A Survey Journal of Visual Communication and Image Representation, 59, 148–159.

[10] Zhang, Y., Wang, J., & Li, S. (2019). Comparative Analysis of Face Detection Algorithms for Video Surveillance Systems. Sensors, 19(14), 3121.

[11] Singh, S.K., Singh, A.K., and Singh, A.K. (2018). A comparative study of face detection algorithms. International Journal of Engineering Research & Technology, 7(4), 2081-2086.

[12] Dong, Y., Zhang, Y., Jiang, H., & Chen, Y. (2020). A Comparative Study of Different Face Detection Algorithms. IEEE Access, 8, 206711-206721.

[13] Zhang, J., Wang, X., & Zhang, Y. (2022). Research on face recognition algorithm based on image processing. PeerJ Computer Science, 8, e8956. [DOI: 10.1371/journal.pmc.8956407]

[14] Ashu Kumar, Amandeep Kaur, and Munish Kumar. Face detection techniques: a review. Artificial Intelligence Review, 52:927–948, 2019.

[15] Sudipto Mondal, Indraneel Mukhopadhyay, and Supreme Datta. Review and Comparison of Face Detection Techniques, pages 3–14. 01 2020.

[16] Paul Viola and Michael J Jones. Robust real-time face detec- tion. International journal of computer vision, 57:137–154, 2004.

[17] Hyung-Ji Lee, Wan-Su Lee, and Jae-Ho Chung. Face recog- nition using fisherface algorithm and elastic graph matching. In Proceedings 2001 International Conference on Image Pro- cessing (Cat. No. 01CH37205), volume 1, pages 998–1001. IEEE, 2001.

[18] Matti Pietika¨inen, Abdenour Hadid, Guoying Zhao, and Timo Ahonen. Computer vision using local binary patterns, vol- ume 40. Springer Science & Business Media, 2011.

[19] Timo Ahonen, Abdenour Hadid, and Matti Pietika¨inen. Face recognition with local binary patterns. In Computer Vision- ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part I 8, pages 469–481. Springer, 2004.