International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 15 • Issue 1 pp. 31-38 Sept 2023-March 2024 www.csjournals.com

Challenges and Review of Biometric Face Detection Technology

Chhavi[1], Kapil Dev[2]
[1]Student, [2]Asst. Professor,
Department of CSE, JMIT, Radaur, Haryana.

Abstract: The method of determining whether or not there are any faces in a given image is known as face detection. Face detection is a popular topic in computer vision, and it's very important in specific applications like surveillance systems. The major goal of this research is to study and summarise the techniques of face detection and find the potential challenges. There are various face detection algorithms but most prominent are CNN, Cascade Classifier, Dlib and facenet. Objective of each method is to find the face in a digital image or frame accurately and timely. There are issues and challenges in finding the faces which are discussed in this research.

I INTRODUCTION

An image/picture is a group of pixels internally stored in the form of bits and bytes. Analyzing the pixel data (binary information) of image and detecting the human face from it is a difficult task. Paul Viola and Michael Jones [1] offered the first authentic and effective research in their paper "Rapid Object Detection by Using a Boosted Cascade of Simple Features" in 2001. It was built on the HAAR theory technique.

II LIBRARY STUDY

The literature review spans the years 2001 through 2021 for the purposes of this dissertation. Face detection has been intensively explored as a key topic in computer vision. According to both the regeneration of the Convolutional Neural Network, face identification is conducted using only a number of machine learning approaches (CNN).

CLASSICAL CONCEPTS OF FACE DETECTION

Viola and Jones [7] describe a face detection framework that can analyze photographs extremely quickly at the same time exceptionally high detection rates were achieved. Three major contributions have been made. The first is the debut of a brand-new picture illustration called a "Integral Image." This allows us to quickly compute the functions of our detector. The 1/3 contribution is a way of "cascading" classifiers, allowing background parts of the picture to be swiftly removed while more computation is focused on possible face-like areas. Overall, the performance of the face detection system is equivalent to that of prior best-in-class systems. This is the first research to use Adaboost with a Haar-like characteristic for training an cascades version that come across face in real time. However, a challenge of those techniques is that their use of susceptible functions (Haar-like functions)

D. McAllester, P. F. Felzenszwalb, and R. B. Girshick [8] they describe a well known approach for constructing cascade classifiers from component-based deformable models along with pictorial systems. They attention ordinarily at the case of star-based models and display how a easy set of rules primarily based totally on partial speculation pruning can accelerate item detection through multiple order of importance without sacrificing detection accuracy. In their set of rules, partial hypotheses are pruned with a series of thresholds. In analogy to possibly about correct (PAC) learning, they introduce the belief of possibly about admissible (PAA) thresholds. Such thresholds offer theoretical ensures at the overall performance of the cascade approach and may be computed from a small pattern of high quality examples. They define a cascade detection set of rules for a well known magnificence of models described through grammar formalism. This magnificence consists of now no longer best tree-dependent pictorial systems however additionally richer fashions which can constitute every element recursively as a combination of different parts.

Akshay Tripathi, S.Murugaveni, Mrinalini Khanna, Aditya Vikram Bhattacharya [9] At the time of taking attendance, the project goals toward the assistance of instructors. Face detection and identification are the primary functions of the device. The project's fundamental principle is face popularity in conjunction with a database backend. Here are preserved the data of the students who attend the magnificence. The type of time stamps used on the server determines the overall attendance. The time stamp provides for the recording of both the hour and the number of individuals that attended the performance. With a time stamp, there are certain exceptions within the grandeur could be included to accommodate students quitting the magnificence or looking to bunk the magnificence. If comparable exceptions occur within the time stamp, the device may be subject to further development. All of the students' questions or concerns could be addressed via the device during a dialogue with the administrator.

TWO-STAGE DETECTORS

T. Gevers, A. W. Smeulders, R. Uijlings, and K. E. van de Sande (R. Uijlings [11] they provided a technique for the localization of objects in an photograph which technique is neuronal and has steps. In the primary step, a difficult localization is carried out through providing every pixel with its neighborhood to a neural net which is ready to signify whether or not this pixel and its neighborhood are the photograph of the search item. This first clear out does now no longer discriminate for function. From its result, regions which would possibly comprise an photo of the item may be selected. In the second one step, those regions are provided to every other neural internet that may decide the precise function of the item in every area. This set of rules is carried out to the trouble of localizing faces in images.

J. Donahue, J. Malik, T. Darrell, and R. Girshick, [12]They present a simple and scalable detection set of criteria that boosts suggest common precision (MAP) by more than 30% when compared to previous best findings. Their method includes the following fundamental insights: To localize and phase objects, Bottom-up area suggestions will be aided by high-capability convolutional neural networks (CNNs). Despite the lack of categorized training material, the combination of guided well before for such an additional work and domain-specific fine-tuning improves actual quality significantly.

ONE-STAGE DETECTORS

M. Mathieu, P. Sermanet, R. Fergus Y. LeCun, D. Eigen, and X. Zhang, [13] they offer an integrated framework for the categorization of Convolutional Networks, localization and detection. They display how a multi scale and sliding window technique may be successfully carried out inside a ConvNet. They additionally introduce a unique deep getting to know technique to localization through getting to know to are expecting object boundaries. Bounding boxes are then collected in preference to suppressed to be able to growth detection confidence. They display that unique responsibilities may be discovered concurrently the use of a single shared network.

Limitations: Traditional and one-level object detection approaches, such as boosted detectors and DPMs, as well as more recent methods, such as SSD [13], during training you will come across a large magnitude mismatch. Except for a small number of them, such detector compares 104-105 possible sites to the images. This imbalance is the source of the following issues:

1. Training seems to be effective since the majority of locations are smooth negatives, implying that contributions do not require entire income signals.
2. Smooth negatives can overpower training, resulting in models that degenerate.

MISCELLANEOUS CONCEPTS

Gang Hua, Zhe Lin, Jonathan Brandt XiaohuiShen, and Haoxiang Li, [14] proposed a complicated discriminative version to as it should be differentiating faces from the backgrounds. They proposed a cascade structure constructed on convolutional neural networks (CNNs) with a high level of discriminative capacity The proposed CNN cascade operates at various resolutions while keeping significant improvement, instantly rejecting surrounding areas within quick short levels of decision making and exhaustively analyzing the small percentage very challenging applications in the involve high selection phase.

Y. Yang, Y. Deng, Y. Yu, and L. Huang [15] Dense Box is an end-to-end integration FCN (completely convolutional neural network) system that predicts bounding containers and object magnificence confidences across all image scales and locations. They proved that if a single FCN is correctly constructed and tuned, it can find multiple exceptional objects in a timely and efficient manner. Second, they show that Dense Box enhances object detection accuracy when combined with landmark localization via multi-mission learning.

Kaiming He Xiangyu Zhang ShaoqingRenJianSun [16] Although deeper neural networks are more harder to teach, they give a residual learning framework that makes it simpler to train networks that are considerably deeper than previously used networks. Rather than acquiring unreferenced skills, In relation to the layer inputs, they specifically renamed the layers as learning residual capabilities. They give comprehensive empirical proof that residual networks are easier to tune and that considerably greater depth can improve accuracy.

Jianfeng Wang, Ye Yuan, Gang Yu [17] with the advancement of the convolutional neural community, the overall performance of face detection has improved significantly. However, occlusion caused by masks and sunglasses continues to be a problematic issue. The worry of excessive false positives is usually associated with the development at the do not forget of those occluded situations. They introduce Face Attention Network (FAN), a new face detector that has the potential to greatly enhance the remembering of the face detection issue in blocked situations. They recommend a novel anchor-degree attention technique that allows you to highlight functions from the face region. Their anchor assigns approach and information augmentation tactics are integrated with their anchor assign approach.

Amit Kumar, AzadehAlavi and Rama Chellappa[18] Along with face modeling, popularity, and verification, keypoint detection is one of the most crucial pre-processing steps for responsibilities. They present an adaptive approach that uses Key Point Estimation and Pose Prediction of Unconstrained Faces Using Learning Efficient H-CNN Regresses to deal with the challenge of face alignment (KEPLER). In recent state-of-the-art methodologies, Convolutional Neural Networks have improved face key factor detection (CNNs). They introduce the H-CNN (Heat map-CNN) framework, which captures well-known worldwide and nearby functions and hence supports accurate key factor discovery. H-CNN is simultaneously taught on facial visibility, fiduciary, and 3-d-pose. The error decreases as the iterations proceed, resulting in modest gradients, necessitating green schooling of DCNNs to mitigate this. For the first four iterations, KEPLER uses international adjustments in posture and fiducially, followed by nearby corrections in the next stage. KEPLER also accurately provides 3-d posture (pitch, yaw, and roll) of the face as a by way of-product.

Deepali G. Ganakwar Vipulsangarm K. Kadam [19] with extensive style of growth in photo and video database, the call for increases for automated exam of this database as it's far bulky in guide information and exam. This paper affords quick insights into a few of renowned and commonly commonplace Techniques of face detection. Face detection approach may be virtually described as a era utilized by pc machine that detects one or numerous human faces ensuing in virtual photo. Recognizing and monitoring the face, estimating pose and expressions, evaluation of face and detecting some other functions of face are the stairs protected in face detection technique. Nowadays, face detection strategies owes one of the maximum lively research regions of computer vision .Considering the face as an item that grabs infinite programs in photo processing makes it difficult project in computer vision. This paper affords a survey of current literature on human face detection machine. Three generally used techniques had been taken into consideration for comparative evaluation on this paper.

Zhanpeng Zhang, Zhifeng Li, Kaipeng Zhang [20] Because of the various postures, illuminations, and occlusions, face identification and alignment in unconstrained environments is difficult. According to recent study, deep learning methods can improve overall performance on such duties. We propose a deep cascaded multi-project architecture in this research, which takes use of the natural correlation among them to improve overall performance. A Researcher present novel idea on-line challenging pattern mining strategy that can improve overall performance mechanically without relying on guide pattern selection as part of the learning process. At the difficult FDDB and WIDER FACE benchmarks for face identification, as well as the AFLW standard for face alignment, our technique outperforms current strategies. While maintaining overall performances

Renad Alharthi, Wafaa Alsubaie, Reem Alshammari, Dana Alqahtani, Rawan Al ramadan, Raneem Alghamdi, Leena Alqarni, Rawan Alsubaie, and Jana Alghamdi [5] The Facial Detection System (FRS) has become a computer-based system using a variety of algorithms to recognize faces that become aware of the human face in virtual photographs, become aware of the individual after which confirm the captured photos via way of means of evaluating them with the facial photos saved with inside the dataset.

Subramanya, Kesava Jayendra Varma, Venkata Sai Harish A and Dr Praveen Kumar S [21] despite the fact that there are numerous techniques to facial detection models, we frequently encounter single facial detection systems from an image. However, in an ever-changing environment, detecting and detecting a single face from an image is not particularly practical. For this, we'll need systems that can detect and recognize many faces from a single image, allowing us to address more real-world problems with fewer photographs. As a result, we introduced HPMR (High Performance Many-face Detection), an improved and efficient model for detecting and recognizing multiple faces from a single image. To recognize faces in this paper, we used Dlib's ResNet network with 29 convolution layers. To estimate face landmarks, the network supports both the Predictor 5 and Predictor 68 models.

Hoai Nam Vu1 Mai Huong Nguyen2 Cuong Pham1 [10] Due to its practicality and convenience of use, face detection is one of most extensively used biometric identification systems. The COVID-19 epidemic has recently spread rapidly over the world, causing major negative consequences. Using face mask in crowded places to prevent the spread of infections has a positive impact on people's health and economic well-being. Masked face detection, on the other hand, is difficult due to the lack of facial feature information. In this study, we propose an approach that combines deep learning with Local Binary Pattern (LBP) features. The masked face was recognized using RetinaFace, a hybrid extra-supervised and self-supervised multi-task learning face. It can deal with a wide range of face scales as a quick but effective encoder. A local binary pattern is also extracted. To make a new face, combine features from the masked face's eye, forehead, and eyebrow sections with features learned from RetinaFace. The development of a consistent framework for recognizing masked faces has been completed.

Agrawal & Samson, [1] Face detection can be accomplished in a variety of ways, one of which is feature extraction, in which the algorithm examines the image directly. For features of the human face that are unique to it. The False Acceptance Rate (FAR) and False Rejection Rate (FRR) are two forms of false acceptance rates discussed in this work (FRR). The likelihood of a system incorrectly recognizing individuals is FAR, while the likelihood of a system failing to identify everyone is FRR (also known as Error Rate)

Ding & Tao [2] Face detection abilities are divided into two categories: frontal and distant. Frontal face detection and pose invariant Face detection is a frontal kind that has been intensively explored and has gradually matured in recent decades thanks to new technologies and methodologies, and it is only concerned with the perspective of the face. Ding and Tao's pose invariant facial detection is a crucial step toward reaching the full potential of facial detection in real-world applications. The article research addresses three degrees of freedom of facial posture change: yaw, pitch, and roll; it covers existing methods that researchers utilise to handle this specific difficulty in field official detection.

Artiges, Caron, Ekenel, Grm, & Struc, [3] discusses the advantages and disadvantages of using CNN for facial detection, particularly in pictures. The quality is poor. This article focuses on the different ways in which an image might be identified and rendered of low quality, and then sends the photos to three CNN models that have already been trained: VGG-Face, GoogLeNet, and SqueezeNet. Blur, contrast/brightness, partial facial occlusion, and noise are all utilized to impair image quality. The study came to the conclusion that blurring was the most difficult feature to cope with in low-quality photographs. The CNN model, as well as a deep learning model, can be taught and developed with the correct architecture and training protocols for detecting faces in low-quality photographs.

Hart, Prikner, & Hartova [4] consider the impact of lighting on faces, particularly shadows, and how they affect accuracy and reliability, particularly for face readers. Biometrics in commercial use. This study involved detecting a face fixed in a position by varying the brightness of a fixed light source surrounding the detector. Lighting has an important impact in facial detection, according to the study's findings.

Boyko, Basytiuk, & Shakhovska [6] Researchers use Dlib and OpenCV to look at one of the most common challenges in computer vision, namely face detection. This article only tries to use HOG in depth when making comparisons and evaluations. The capacity to recognize faces with accuracy is merely a comparison of the length of time it takes to discover faces in a series of photographs when compared to the computer. Face recognition software looks for patterns in the shape of a person's face.

III POTENTIAL ISSUES OR CHALLENGES

Growing interest in face detection is good, but it also proves to be a difficult undertaking when it comes to issues that have consistently hampered its quality of service. These are the problems which create situations that are uncooperative, and give a large number of facial appearances and expressions.

Illumination

A light variation in an image is called illumination. As shown in figure 1 A small variation in lighting gives a big problem in finding the face and can have a major impact on the results. Capture the picture with lighting change, the same person in a nearly same facial expression and pose, the results will be greatly different. The appearance of the face is radically altered by illumination. We can see in various examples that the difference between two identical faces clicked with different lighting is more than the difference between two different faces clicked with the same lighting.

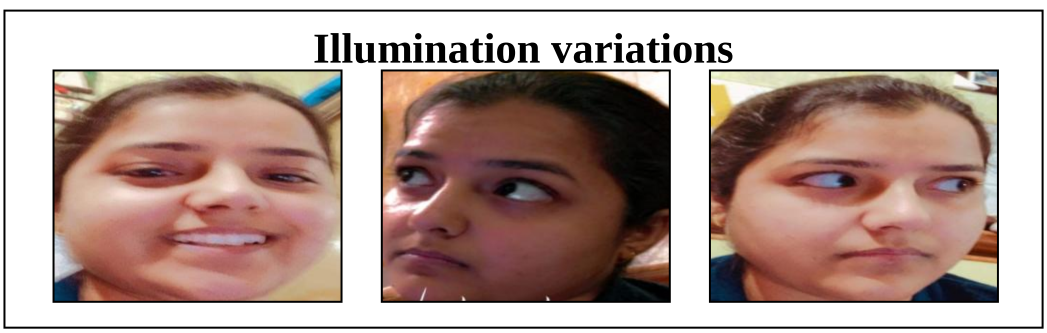

Figure 1: Illumination Variation

Pose

The change in the pose of a person is very sensitive in face detection methods. As shown in figure 2 when a person's head moves in any direction or the change in angle of viewing, it is called that the pose has been changed. The effect of head motions or different camera POVs generally induce similarity variances, causing Face detection results to plummet. When the rotation angle is increased, identifying the true face becomes more difficult. If the database only has a frontal image of the face, the detection may be inaccurate or non-existent.

Figure 2: Pose Variations

Occlusion

Occlusion means some areas of the face are overlapped, which prevents the whole face used as an input image. Handling occlusion is most difficult situation in a face identification method. It is common in real-world scenarios and is caused by beards, moustaches, and accessories (goggles, caps, masks, and so on). The existence of such components diversifies the subject, making computerized face detection a difficult nut to crack.

Expressions

As shown in figure 3 Next point to consider is the different face expressions of the person. Expressions such as happiness, sadness, rage, fear, and surprise, are some examples of human face expressions. Some Micro-expressions of human face that reveal fast facial changes. One's emotional state appears on its face, and efficient identification becomes difficult in the wake of such emotions, which are numerous.

Figure 3: Expression Variations

Low Resolution

Finding faces or other objects in low resolution images is very difficult problem. Small scale cameras, such as CCTV cameras, ATM cameras etc. can provide these low resolution images. Because the camera is not very close to the face, these cameras can only catch a 16*16 facial region. Because the most of the pixels in such a low-quality image are lost, it does not communicate much information. Identifying people's faces is a difficult task.

Ageing

The appearance/texture of a person's face varies with time and reflects their age, making facial recognition systems more difficult to use. As people become older, their facial characteristics, shapes/lines, and other features change. It's employed for picture retrieval and long-term visual monitoring. For accuracy verification, the dataset for a separate age group of people over a period of time is calculated. The identification procedure is based on feature extraction, which includes wrinkles, blemishes, brows, hairstyles, and other basic traits.

Model Complexity

Current framework facial recognition systems rely on a complex Convolutional Neural Network (CNN) architecture that is "too deep" for real-time performance on embedded devices. An ideal face detection system should be able to handle variations in illumination, emotion, position, and occlusion. It should be scalable to a large number of users with low image capturing requirements during registration while avoiding complex design.

IV Conclusion

The face is the most visible and important feature of a person and its unique characteristics make it essential for human identity. Various techniques and technologies are used around the world to improve the accuracy and reliability of face detection. Healthcare, security, defense, forensics, and transportation are all areas where this ever-expanding technology is being used, and more accuracy is required. However, while developing face detection technology, some obstacles are universal, such as position, occlusion, expressions, ageing, and so on. In the world of computer vision, face detection remains a difficult subject. It has attracted a lot of interest in recent years due to its numerous uses in various fields. Despite the fact that there is a lot of research going on in this area, face detection algorithms are far from perfect in terms of performing well in all actual conditions. A brief review of challenges, methods, and applications in the field of face detection was provided in this paper. There is still more efforts to be done in order to develop approaches that represent how humans recognize faces and make the most use of the face's time development for detection.

Reference

1. Agrawal, A., & Samson, S. (2016).A Review on Feature Extraction Techniques and General Approach for Face Detection. International Journal of Computer Applications Technology and Research, 5(3), 156-158.
2. Ding, C. & Tao, D. (2016). A Comprehensive Survey on Pose-Invariant Face Detection. ACM Transactions on Intelligent Systems and Techzeroogy, Apr (7), 1–40.
3. Artiges, A. Caron, M. Etonel, H .K Grm, K., & Struc, V. (2017). Strengths and Weaknesses of Deep Learning Models for Face Detection against Image Degradations. IET Biometrics, 7(1), 81-89.
4. Hart, J., Prikner, P., & Hartova, V. (2018). Influence of Face Lighting on the Reliability of Biometric Facial Readers .Agronomy Research, 16, 1025–1031.
5. Alghamdi, J., Alharthi, R., Alghamdi, R., Alsubaie, W., Alsubaie, R., Alqahtani, D. Alshammari, R. (2020). A Survey on Face Recognition Algorithms. 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS).
6. Boyko, N, Basytiuk, O., & Shakhovska, N. (2018). Performance Evaluation and Comparison of Software for Face Detection, based on Dlib and Opencv Library.2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)(AUG), 478–482.
7. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Detection, 2001.CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, volume 1, pages I–I. IEEE, 2001

8. P.F .Felzenszwalb ,R.B.Girshick, and D. McAllester. Cascade object detection with deformable part models, in CVPR, 2010
9. Class Monitoring System Tools MTCNN and Haarcascade Classifier 2018 Aditya Vikram Bhattacharya1, Mrinalini Khanna2, Akshay Tripathi3, S.Murugaveni ,Dept. of Information and Telecommunication, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur, Kancheepuram District, Tamil Nadu, India
10. Masked face detection with convolutional neural networks and local binary patterns Hoai Nam Vu1 · Mai Huong Nguyen2 Cuong Pham1 Accepted: 26 July 2021 The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021
11. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders .Selective search for object detection .IJCV, 2013.
12. R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, 2014.
13. P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat: Integrated detection, localization and detection using X Convolutional networks, in ICLR, 2014
14. H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua. A convolutional neural network cascade for face detection .In Proceedings of the IEEE Conference on Computer Vision and Pattern Detection, pages 5325–5334, 2015.
15. L .Huang, Y. Yang, Y. Deng, and Y. Yu .Dense box: Unifying landmark localization with end to end object detection. arXiv preprint arXiv:1509.04874,2015.
16. Kaiming He, Xiangyu Zhang, ShaoqingRen and Jian Sun: Deep residual learning for image detection .arXiv preprint arXiv:1512.03385, 2015
17. Jianfeng Wang, Ye Yuan, Gang Yu Face Attention Network: An Effective Face Detector for the Occluded Faces arXiv:1711.07246v2 [cs.CV] 22 Nov 2017
18. Amit Kumar, Azadeh Alavi and Rama Chellappa, KEPLER: Keypoint and Pose Estimation of Unconstrained Faces by Learning Efficient H-CNN Regressors arXiv:1702.05085v1 [cs.CV] 16 Feb 2017
19. Comparative Analysis of Various Face Detection Methods2019 Deepali G. Ganakwar Dr. Babasaheb Ambedkar Marathwada University, Dr .Vipulsangram.K..Kadam
20. Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Senior Member, IEEE, and Yu Qiao, Senior Member, IEEE2016
21. dr praveen kumar s 1, kesava jayendra varma v2, subramanya v3, venkata sai harish a 4 a multiple face detection system with dlib resnet network using deep metric learning Journal of Critical Reviews ISSN- 2394-5125 Vol 7, Issue 6, 2020