ISSN 2320 - 2602
Volume 13 No. 1, January 2024
International Journal of Advances in Computer Science and Technology
Available Online at http://www.warse.org/IJACST/static/pdf/file/ijacst071312024.pdf
https://doi.org/10.30534/ijacst/2024/071312024

Review: Tiny Face Detection and Recognition Techniques

Akshatha H [1], Ananya Preethi[2], Spoorthi H S[3], Srinidhi C K[4], Tejas Ravi[5], Rizwan shaikh[6]
Students, Department of Computer Science and Engineering[12345]
Assistant Professor, Department of Computer Science and Engineering
[1]Alva's Institute of Engineering and Technology, India, akshatahmd@gmail.com
[2]Alva's Institute of Engineering and Technology, India, ananyapaul369@gmail.com
[3]Alva's Institute of Engineering and Technology, India, spoorthihs13@gmail.com
[4]Alva's Institute of Engineering and Technology, India, nidhick28042002@gmail.com
[5]Alva's Institute of Engineering and Technology, India, tejasravi64@gmail.com
[6]Alva's Institute of Engineering and Technology, India, rizwanshaikh@aiet.org.in

Received Date : November 25 , 2023 Accepted Date : January 3, 2024 Published Date : January 07, 2024

ABSTRACT

The exponential growth of video and image databases has created a need for intelligent systems to automatically analyze information, as manual efforts are no longer feasible. Faces plays a crucial role in social interaction, conveying identity and emotions, thus requiring efficient and accurate analysis. Deep learning techniques has brought about a significant revolution in face detection, despite their increased computational requirements.
This paper presents a comprehensive analysis of representative deep learning-based methods for face detection, focusing on their accuracy and efficiency. It also compares and discusses popular and challenging datasets, including their evaluation metrics. Additionally, a thorough comparison of successful deep learning-based face detectors is conducted, evaluating their efficiency using Floating Point Operations (FLOPs) and latency as metrics.
The results and findings of this study can serve as a valuable guide for selecting suitable face detectors for various applications. Moreover, they can contribute to the development of more efficient and accurate detectors. The paper aims to address the pressing needs for intelligent systems that can automatically understand and analyze visual information in an increasingly data-driven world.

Key words; Deep learning based face detection, Facial expression recognition, Neural Network Architectures

1. INTRODUCTION

Face detection poses a significant challenge in the field of object recognition, particularly when it comes to detecting small objects. In this context of face detection, we delve into three crucial aspects: scale invariance, image resolution, and contextual reasoning. Scale invariance is a fundamental property in recognition and object detection systems, enabling robust detection across different scales[6]. The exponential growth of video and image databases has made it increasingly burdensome for humans to manually process and analyze the vast amount of visual data available. Faces hold paramount importance in social interactions, serving as a means to convey identity and emotions. To address the demanding task of automatically comprehending and examining visual information, deep learning technique has emerged a powerful solution, exhibiting remarkable breakthroughs in face detection. However, these advancements come at the cost of heightened computational requirements[8].
This paper aims to provide a comprehensive exploration of deep learning-based methods for face detection, with a specific focus on their accuracy and efficiency. The goal is to shed light on the advancements in this field, analyzing various techniques and their performance. Additionally, the paper highlights the need to strike balance between accuracy and computational efficiency on face detection systems.
Figure 1 showcases examples of successful face detection, demonstrating the capabilities of the discussed methods[13].

Figure 1: Examples of face detection

2. DEEP LEARNING-BASED FACE DETECTION METHODS

The paper provides an overview of representative deep learning-based methods employed for face detection. These methods leverage the power of deep neural networks to extract discriminative [5] features from images and make accurate predictions. Architectures such as Convolutional Neural Networks (CNNs), Region-based Convolutional Neural Networks (R-CNNs), and Single Shot MultiBox Detectors (SSDs) are among the popular choices in the field. Each method is described in detail, highlighting its strengths and weaknesses in the terms of accuracy and computational complexity[9].

3. FACE DETECTION AND RECOGNITION

Face detection is a critical component of facial analysis systems and can be several steps:

A. Input: An image is provided as input to the face detection system. The image can have different formats, sizes, resolutions, and may even consist of video frames.

B. Pre-process: The image undergoes pre-processing to eliminate unwanted noise caused by lighting and the environment. This step also involves normalizing the image to ensure consistent processing.

C. Classifier: The classifier plays a crucial role that determining whether the image contains a face or belongs to the non-face class. The classifier utilizes information learned during training to make this decision.

D. Output: The output of the face detection system indicates whether the original input image contains a face or not. Some systems also provide information, such that location of the detected face within the image [3].

Various features that face have been studied in the context of face detection and recognition:

1. **Position:** Detecting the precise position of face within a image is crucial. Face detectors aim to accurately identify the spatial location of a face, regardless of its position within the image.
2. **Scale:** Faces can appear at various scales within an image due to differences in distance from the camera or image resolution. Face detection algorithms need to handle faces of different sizes to ensure robust detection.
3. **Orientation:** Faces can be oriented in different directions, such as frontal, profile, or tilted angles. Face detectors should be capable of detecting faces regardless of their orientation.
4. **Illumination:** Lighting conditions can vary significantly in images, leading to a variation in face appearance. Effective face detect algorithms are robust to changes in illumination to ensure accurate detection under various lighting conditions[7].

Face detection can be a challenging task due to several factors. Faces exhibit variations in terms like age, skin color, facial expressions, making it difficult to establish a universal face model. Additionally, lighting conditions, background geometries and image qualities further complicate the detection process. The idea face detector should be capable of detecting faces reliably under any lighting conditions and against any background [1].

By leveraging advanced techniques in computer vision and machine learning, researchers strive to develop face detection algorithms that can handle these challenges and reliably detect faces in diverse scenarios.

4. PROPOSED METHOD

In the proposed method, as depicted in Figure 2, two processing stages are involved. The first stage focuses on increasing image resolution, while the second stage revolves around recognizing the face image. The primary objective of the initial stage is to enhance image resolution, and it is achieved through the utilization of a Convolutional Neural Network (CNN) as proposed [4]. On the other hand, the subsequent stage emphasizes classification and employs three fully connected layers. The computations in the CNN align with those presented, while calculations in the fully connected (FC) layer employ Multilayer Neural Networks. For reference, the system architecture is observed in Figure 2.

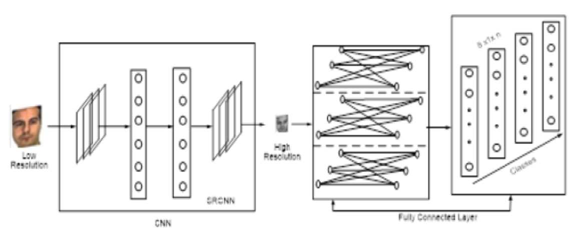

Figure 2: The framework of VsFRCNN

4. OpenCV

OpenCV primarily focused on image processing. The focus is on image processing and can be implemented on the algorithms. It is a BSD license to be used in projects. OpenCV includes Face Detection. The face detection can import:

1. sys – perform a console related functions that like taking a input from a console.

2. cv2 – OpenCV object or module that has the functions of image processing.

Haar cascade Object Detection Face & Eye OpenCV Code

```
import numpy as np
import cv2

# multiple cascades:
https://github.com/Itseez/opencv/tree/master/data/haarcascad
es
#https://github.com/Itseez/opencv/blob/master/data/haarcasca
des/haarcascade_frontalface_default.xml
face_cascade =
cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
#https://github.com/Itseez/opencv/blob/master/data/haarcasca
des/haarcascade_eye.xml
eye_cascade = cv2.CascadeClassifier('haarcascade_eye.xml')

cap = cv2.VideoCapture(0)

while 1:
  ret, img = cap.read()
  gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
  faces = face_cascade.detectMultiScale(gray, 1.3, 5)

  for (x,y,w,h) in faces:
    cv2.rectangle(img,(x,y),(x+w,y+h),(255,0,0),2)

    roi_gray = gray[y:y+h, x:x+w]
    roi_color = img[y:y+h, x:x+w]

    eyes = eye_cascade.detectMultiScale(roi_gray)
    for (ex,ey,ew,eh) in eyes:

cv2.rectangle(roi_color,(ex,ey),(ex+ew,ey+eh),(0,255,0),2)

  cv2.imshow('img',img)
  k = cv2.waitKey(30) & 0xff
  if k == 27:
    break

cap.release()
cv2.destroyAllWindows()
```

5. DEEP BELIEF OF NETWORK

Deep Belief of Network is a deep learning architecture is widely used in machine learning. It is considered an graphical model that incorporates a probabilistic approach to reconstruct its inputs. DBN consists of multiple layers and has directed connections between them, as depicted in Figure 3.

In the DBN architecture, each layer is composed of a simple and unsupervised network such that Boltzmann machines, autoencoders. The hidden layer of each subnetwork serves the visible layer for a subsequent layer, creating a hierarchical structure. This layer-wise learning enables the DBN to capture intricate patterns and representations in the input data[2].

Figure 3 provides a schematic overview that DBN architecture, illustrating directed connections between layers. The arrows depict the flow of information in a graphical model.

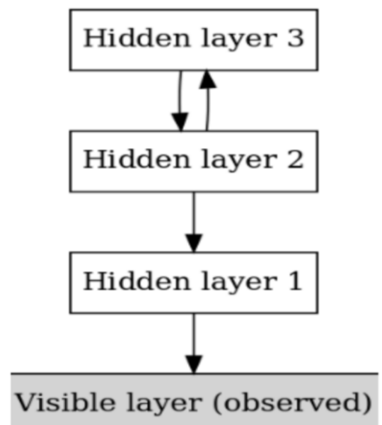

Figure 3: DBN architecture

6. SPEED FOCUSING FACE DETECTORS

Table 1: State-of-the-art Open-Source Models for Real-Time Face Detection

Model	AVG TFLOPs	AVG Latency (ms)		
		Forward (GPU)	Forward (CPU)	Post-Proc
RetinaFace	0.201	131.60	809.24	8.74 (GPU)
CSP	0.579	154.55	1955.20	27.74 (CPU)
SRN	1.138	204.77	2933.16	8.71 (GPU)
DSFD	1.559	219.63	3671.46	76.32 (CPU)

We conducted tests on various open-source face detectors using a 720P image containing multiple faces at a scale of 1.0. The aim of these face detectors is to achieve real-time face detection while keeping the computational that costs significantly around lower than the models discussed earlier [10].

The average FLOPs (AVG TFLOPs) and latency (AVG Latency) were calculated by running the tests for each model 100 times. The "Post-Proc" stage refers to post-processing steps such as decoding from anchors and non-maximum suppression (NMS), where we utilized the original processing code of each model.

Based on our evaluation, we present Table 1, which highlights the network architectures, Average Precision (AP), FLOPs, and efficiency of the most popular speed-focusing face detectors collected from github.com. These models exhibit computational costs in the range of GFLOPs or 10 GFLOPs, significantly lower than the earlier discussed models[12].

The review of these speed-focusing face detectors provides a valuable insight in their network architectures, performance in terms, computational costs (FLOPs), and overall efficiency. This information helps in selecting suitable face detection models for real-time applications, taking into a account both accuracy for computational requirements [11].

The Neural Network Approach, inspired by a information processing mechanisms of biological nervous systems, particularly the brain, is an effective paradigm for information processing. It revolves around a unique structure of interconnected processing elements called neurons, working collaboratively to solve specific problems. This approach will be based on the concept of learning by example, similar to how humans learn from experience.

(ANNs) is configured for pattern recognition and data classification. The learning process involves adjusting the synaptic connections between the neurons, analogous to the synaptic connections in biological systems.

ANNs consist that large number of interconnections processing at elements (neurons) that work together to process information and make predictions. The structure of the network, including the number of layers and the connections between neurons, can vary depending on the specific application and problem at hand.

By leveraging the Neural Network Approach, ANNs have shown remarkable capabilities in various domains, including image and speech recognition, natural language processing, and more. Their ability to learn from data and adapt to complex patterns makes them powerful tools for solving challenging problems[14].

5. CONCLUSION

This paper presents a comprehensive exploration of deep learning-based methods for face detection, offering insights into their accuracy and efficiency. By comparing and analyzing popular datasets, evaluation metrics, and successful face detection models, it provides valuable guidance for selecting suitable detectors for various applications. Furthermore, the paper has highlights that has importance of developing more efficient and accurate face detectors address the growing need for automatic understanding and examination of visual information.

REFERENCES

[1] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, “The Pascal Visual Object Classes (VOC) Challenge,” International and Journal of Computing Vision, vol. 88, no. 2, pp. 303–338, 2010.

[2] S. Yang, P. Luo, C.-C. Loy, and X. Tang, “WIDER Face: A Face Detection Benchmark”, Proceedings of IEEE International Conference on a Computer Vision of Pattern Recognition, 2016.

[3] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, “Face Detection Based on Multi-Block LBP Representation,” in International Conference on Biometrics, 2007.

[4] Face Detection, Sameer Hashmi.

[5] Discriminant and Analysis for Recognize the Human Face Image, Peiyun Hu

[6] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, “Face Detection Based on Multi-Block LBP Representation,” in International Conference on Biometrics, 2007.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[8] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, “Face Recognition: A Literature Survey,” Technical Report CART-TR948, University of Maryland, Aug. 2002.

[9] Finding Tiny Faces by Ramana “Training Support Vector Machines: An Application to Face Detection,” .

[10] P. Viola, M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,”.

[11] M. Turk and A. Pentland (Jun. 1991). “Face Recognition Using Eigenfaces.” CVPR IEEE Computer Society.

[12] Akmalbek, A.; Djurayev, A. Robust shadow removal technique for improving image enhancement based on segmentation method. *Commun. Eng.* **2016**, *11*, 17–21.

[13] Lawrence S, Giles C L, Tsoi A C, et al. Face recognition: [J]. IEEE Transactions on Neural Networks, 1997, 8(1):98-113.

[14] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, “Retinaface: Single-stage dense face localisation in the wild,” ArXiv, vol. abs/1905.00641, 2019