Review

JOURNAL OF UNIVERSITY OF BABYLON

For Pure and Applied Sciences (JUBPAS)

Vol.32; No.2.| 2024

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

A Face Detection System: A Comprehensive Survey

Faqeda Hassen[1*], Mohammed Abdullah Naser [2]

1College of Science for Women,University of Babylon, faqida1989@gmail.com,Babylon, Iraq.

2College of Science for Women, University of Babylon, ،wsci.mohammed.abud@uobabylon.edu.iq, Babylon, Iraq.

***Corresponding author email:** mohamed127612@gmail.com ; mobile: 07801273207

نظام اكتشاف الوجه : دراسة شاملة

فاقدة حسن كريم [1*]، محمد عبد الله ناصر [2]

1 "كلية العلوم للبنات، جامعة بابل، faqida1989@gmail.com، بابل، العراق"

2 "كلية العلوم للبنات، جامعة بابل ، wsci.mohammed.abud@uobabylon.edu.iq ، الحلة ، بابل"

Accepted:	15/3/2024		Published:	30/6/2024

ABSTRACT

The domain of facial detection has witnessed significant advancement, mostly attributed to the progress in machine learning and deep learning methodologies. These enhancements have resulted in a fundamental change in the field. Face detection algorithms are extensively utilized in several domains, such as security, surveillance, and social networking applications. This survey has thoroughly investigated various methodologies for facial identification, with a focus on the challenges, applications, and evolution from traditional to deep learning-based approaches. Furthermore, it provides important insights into commonly used datasets for facial detection, including a comprehensive examination of their unique characteristics. The study is structured sequentially, commencing with an introduction to the assessment of significant previous research and the challenges faced by face detection algorithms. Subsequently, the investigation delves into different practical uses, delineates methods for identifying, and scrutinizes frequently employed data sets. The transition from traditional classifiers to deep learning models in face identification technology represents a significant improvement in addressing the complexities of face detection in various environments. Despite the obstacles that now exist, the continuous advancements in deep learning offer promising solutions that enhance the accuracy and efficiency of face identification systems in a wide range of applications.

Keywords: face detection, machine learning, deep learning, spatial domain and CNN.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

1. INTRODUCTION

The detection of faces is a fundamental and significant problem within the field of computer vision. In today's world of multimedia technology, the significance of image and video has reached unimaginable heights. The human face holds considerable importance as a prominent feature in both images and video[1].

Some research [2,3] that focus of the studies has been chiefly directed towards the domain of detection and analysis related to humans, with an emphasis on face aspects. This pertains primarily to their utility in many domains; this encompasses a range of technologies, including biometrics for access control, monitoring platforms, and considerable advancements made in this domain, the task of accurately detecting faces in unpredictable settings continues to be a relatively specific area of research [4]. However, numerous techniques are employed for facial detection. Before performing facial analysis operations, it's crucial to first achieve facial detection. Previous attempts at detection relied on a conventional method involving the extraction of generated features from the image, utilizing multiple classifiers to accurately identify facial areas [5].

The Haar cascade classifier [6] was among the techniques used. However, the approach employed in this study utilized a Histogram of Oriented Gradients (HOG) technique, followed by the application of a Support Vector Machine (SVM) algorithm [7]. Despite these efforts, accurately identifying faces in challenging images with resolution variations remains a limitation within the WIDER FACE facial detection dataset [8]

In recent years, deep learning, intense Convolutional Neural Networks (CNN), has achieved significant success across several computer vision applications [9]. Hence, it can be concluded that face detectors can get much superior detection rates in comparison to conventional cascaded classifiers when employing Faster R-CNN [3], YOLO [4], or single shot detector (SSD) [5]. Furthermore, various types of detectors may be used, such as multi-task cascaded convolutional networks (MTCNN) [5,6].

The detection process, involving manual feature building and sliding window technique, is complex and computationally intensive, resulting in limited detection accuracy. However, the aim of this work is to conduct a comprehensive examination of various methodologies for detecting faces in digital images and to explore the problems and applications associated with face detection. Additionally, the study provides an overview of commonly used datasets for face identification and describes their properties in detail.

The rest of the paper is organized as follows: Section 2 addresses the most important related previous studies, Section 3 presents the challenges in face detection systems. Section 4 discusses different applications of the same topic. While the section 5 shows the techniques used in detecting faces. followed by Section 6, which reviews the most datasets that used the field of research. Finally, the conclusions.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

2. RELATED WORKS

This section provides an overview of several methodologies employed for face detection, spanning from initial approaches to present-day advancements. Several methodologies for detecting and identifying facial features have been developed in recent years.

The research[6] introduces a new approach for cascaded convolutional neural networks with two primary stages. In the initial phase, a low-pixel candidate window is employed as an input, enabling the rapid extraction of the candidate window by the shallow convolutional neural network. During the second stage, the window from the previous stage is resized and utilized as an input to the corresponding network layer. The detection accuracy of the discrete score on the FDDB dataset is 93.4%.

The study[7] presents a comparative analysis of Two-stage and One-stage detection models in the context of face detection tasks. Additionally, it introduces a multi-task convolutional neural network (MTCNN), which obtains an accuracy of 85.7% when applied to the image of a wider face.

The research [8]provides a novel approach for enhancing the performance of face detection algorithms. Resilience by collecting knowledge about minor facial features on complex images, they adopt VGG16 as the underlying convolutional neural network (CNN) architecture. They perform fusion on two specific layers, conv4_3 and conv5_3, following dimension reduction and bilinear up sampling techniques. This fusion process ultimately yields the desired detection feature map and achieves accuracy scores of 95.7%, 94.9%, and 89.7% on the easy, medium, and complex, respectively, on the WIDER FACE dataset.

The paper[9] Proposed a method to develop an architectural framework to improve the Region-based Convolutional Neural Network (Faster R_CNN) architecture by integrating global and local information. The architecture consists of two primary networks, the region proposal network and the second network, which distinguishes between face and non-facial entities; according to the findings of the experiments, the proposed approach has a recall rate of 92.09 %.

in this paper[10] They suggested use of multiscale Hybrid Pyramid Convolutional Network (HPCNet), a one-stage fully convolutional network with three modules: Hybrid Dilated Convolution, Hybrid Feature Pyramid, and Context Information Extractor. The HPCNet also introduces an enhanced Online Hard Example Mining technique for improved face detection accuracy. On the Easy subset of WIDER FACE, the approach achieved an accuracy of 93%; on the Medium subset, it achieved 92%; on the Hard subset, it achieved 84%.

In This paper [11] aims to use YOLO-face, a face detector based on YOLOv3, which uses a more accurate regression loss function and anchor boxes for face detection. The updated method

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

maintains quick detection speed and improves accuracy, surpassing YOLO and its variants in tests on WIDER FACE and FDDB data sets; the accuracy of the approach was 78%, 73%, and 47% for the easy, medium, and complex images, respectively.

This research[12] presents a novel two-level face detection model, SR-YOLOv5, which aims to tackle the challenges of detecting dense small faces in real-world contexts. The study's initial phase involved optimizing the backbone and loss function employed in YOLOv5. Subsequently, efforts were made to enhance face detection in scenarios characterized by blurriness or low-resolution conditions. The approach achieved 96.3%, 94.9%, and 88.2% accuracy for easy, medium, and high difficulty levels, respectively.

The paper[13] demonstrates the existence of the prediction made by target tracking and shows how face detection and target tracking are related. Integrates the tracking results into the tracking algorithm as a temporal feature. Rest-Single Shot Multi-Box Detector(Rest-SSD) face detection reduces the likelihood of the incident. The missed detection problem arises due to changes in position and occlusion of the face. Experimental findings show an accuracy of 93%, 92%, and 83% for the simple, medium, and complex images.

In this paper[14], The Haar Cascade Classifier (HCC) is suitable for real-time face identification. In order to address the issue of face spoofing detection, the research suggests a particular_ Local Binary Pattern (LBP) network. Deep learning is used in the proposed network to combine hand-crafted qualities, and statistical histograms are used to reduce network parameters. The HCC algorithm demonstrates a face detection rate over 90% when applied to photographs with a basic background. Furthermore, it achieves a face detection rate of 93.24% when confronted with images featuring a complicated background.

The objective of this study [14]is to investigate and evaluate Machine Learning (ML) technologies for the purpose of identifying and recognising individuals who are wearing face masks in various forms of media, including pre-recorded movies, images, and real-time scenarios. The objective of the project is to develop a real-time Graphics User Interface (GUI) that incorporates an Automated Facial Recognition system and a Mask Detection system. The techniques employed in the suggested methodology consist of Principal Component Analysis (PCA) and the HAAR Cascade Algorithm.The aforementioned model has a level of accuracy reaching 99%.

This study [15] investigates facial identification in judicial settings using the Viola-Jones method and convolutional neural network. The approach leverages computational efficiency and spatial resource utilization, resulting in a 99.5% accuracy rate when implemented in real-time conditions and under appropriate illumination.

This study [16]. introduces the Retina Net baseline, a one-stage face detector designed to address the complex face detection task. The network enhancements increased the detection process's speed and accuracy. Two well-known datasets, namely WIDER FACE and FDDB, were used in the experiments. In the context of the WIDER FACE benchmark, it was observed that the detection model achieved a notable accuracy rate of 95.6% in successfully determining and localizing

Table (1): Performance Different Face Detection Methods

Researcher(s) Name	The Year	Method (s)	Dataset used	Accuracy
Wankou Yang and et al.[6]	2019	Cascade convolution structure	FDDB datasets	93.4%.
Ning Zhang et al [8].	2020	MTCNN	wider face data set, Pascal VOC database	85.7%
Zhishuai , and et al.[8]	2020	CNN	WIDER FACE dataset.	92%*
Maliki et al.[9]	2020	improve the Faster R-CNN	WIDER Face and FDDB datasets	92.09%
Shaoqi Hou and et al.[10]	2021	Hybrid Dilated Convolution, Hybrid Feature Pyramid, and Context Information Extractor	WIDER FACE	89%*
Chen ,and Weijun[11]	2021	YOLO-face based on YOLOv3,	WIDER FACE and FDDB data sets	66%*
Qingqing Xu and et al[12]	2021	SR-YOLOv5	WIDER FACE	92%*
Yilin Liu, and et al.[13]	2022	Rest-SSD	WIDER FACE and FDDB data sets	89%*
Nandkumar Kulkarni, and et al[14]	2022	HCC, LBP	camera	91%*

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية
ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Achyutha ,and et al [17]	2022	(PCA) and HAAR Cascade	camera	95.6%
Tameem Hameed Obaida , and et al[15]	2022	Viola-Jones method and CNN	webcam	99.5%
Dilnoza Mamieva , and et al.[16]	2023	Retina Net baseline	WIDER FACE and FDDB.	95.6%

***** The average accuracy for images classified as easy, medium, and hard.

Accuracy: refers to the ratio of accurately predicted observations to the total number of observations. Ratio of true positive predictions to the overall number of positive predictions [2].

$$\boldsymbol{Accuracy = \frac{TP + TN}{TP + FP + FN + TN}} \qquad \textbf{[2]}$$

3. CHALLENGES IN FACE DETECTION SYSTEMS

The challenges associated with face detection contribute to reduced accuracy and detection rate. Therefore, studying and identifying the challenges is important and necessary. There are many of them as shown in Figure (1). Below several examples of them:

1. The facial pose of a face image can be influenced by the camera's position, perhaps resulting in a misalignment of the eyes along the horizontal axis [18].
2. Variations in illumination intensity can lead to an uneven distribution of gray levels in facial images, resulting in significant local contrast variations. These variations can have a detrimental impact on the effectiveness of face detection algorithms [7].
3. Detecting face occlusion, which refers to situations where the face is obstructed by items such as scarves, glasses, and helmets, is an essential security measure to address the incidence of criminal activities [14].
4. Distance between a camera and a human face can affect detection accuracy.
5. Face detection faces significant challenges due to facial expressions in images that deviate from the expected norm.
6. The presence of several objects in an image, referred to as a complex background, has a detrimental effect on the accuracy and efficiency of face detection.
7. The image is characterized by many human faces, making it challenging to identify and detect these features accurately [14].

JOURNAL OF UNIVERSITY OF BABYLON
For Pure and Applied Sciences (JUBPAS)
Vol.32; No.2.| 2024

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Figure (1) shows different types of challenges for face detection[20]

4. APPLICATIONS OF FACE DETECTION SYSTEM

The use of face detection systems in many fields and industries has been the subject of considerable interest and research. Below is a sample of these various applications in which face detection systems are applied. The versatility and utility of face detection continues to expand as research and technological advances drive further innovation in this field [12,14].

1. Facial Recognition: Facial recognition is one of the most widely known and important applications of face detection. Once faces are detected, facial recognition algorithms can match the detected faces against a database of known individuals, enabling identity verification, access control systems, and personalized user experiences. Facial recognition has applications in law enforcement, surveillance, banking, and mobile device security.

2. Emotion Analysis: Face detection plays a crucial role in emotion analysis applications. By detecting and analyzing facial expressions, such as smiles, frowns, or raised eyebrows, face detection systems can infer emotions and sentiments. This has applications in market research, customer feedback analysis, human-computer interaction, and personalized digital content delivery.

3. Age and Gender Estimation: Face detection can be used to estimate the age and gender of individuals in images or videos. By analyzing facial features and patterns, algorithms can estimate an individual's approximate age range and determine their gender. This has applications in targeted marketing, content filtering, and demographic analysis.

4. Human-Computer Interaction: Face detection enables natural and intuitive human-computer interaction. By detecting and tracking facial movements and gestures, systems can interpret

user intentions and provide personalized experiences. Applications include gesture-based control systems, virtual reality, augmented reality, and gaming.

5. Surveillance and Security: Face detection plays a vital role in surveillance and security systems. By detecting and tracking faces in real-time, these systems can identify and monitor individuals of interest, detect unauthorized access attempts, and enhance public safety. Face detection is used in airports, public transportation, critical infrastructure, and crowd monitoring.
6. Biometric Authentication: Face detection is a crucial component of biometric authentication systems. By capturing and verifying an individual's unique facial features, face detection enables secure and convenient authentication methods. Applications include unlocking smartphones, accessing secure facilities, and online identity verification.
7. Social Media and Photo Management: Face detection is used extensively in social media platforms and photo management applications. By automatically detecting and tagging faces in photos, these systems enable easy organization and searching of images based on specific individuals. This simplifies photo sharing, enhances user experience, and facilitates social connections.
8. Automotive Safety: Face detection is being utilized in automotive safety systems, such as driver monitoring and drowsiness detection. By monitoring the driver's face and eye movements, these systems can detect signs of fatigue, distraction, or inattentiveness, alerting the driver and preventing accidents.
9. Medical Applications: Face detection technology has extensive applications in the medical domain, playing a crucial role in multiple facets of patient care and treatment. A critical use case is in identifying face abnormalities, where the technology can aid healthcare practitioners in recognizing and diagnosing medical disorders based on facial characteristics.
10. Face detection technology is utilized in military settings and counterterrorism efforts to employ sophisticated surveillance and recognition systems for real-time identification of people. This technology is essential for improving security measures and operating capabilities in regions that are sensitive or have a high level of risk.

5. FACE DETECTION TECHNIQUES

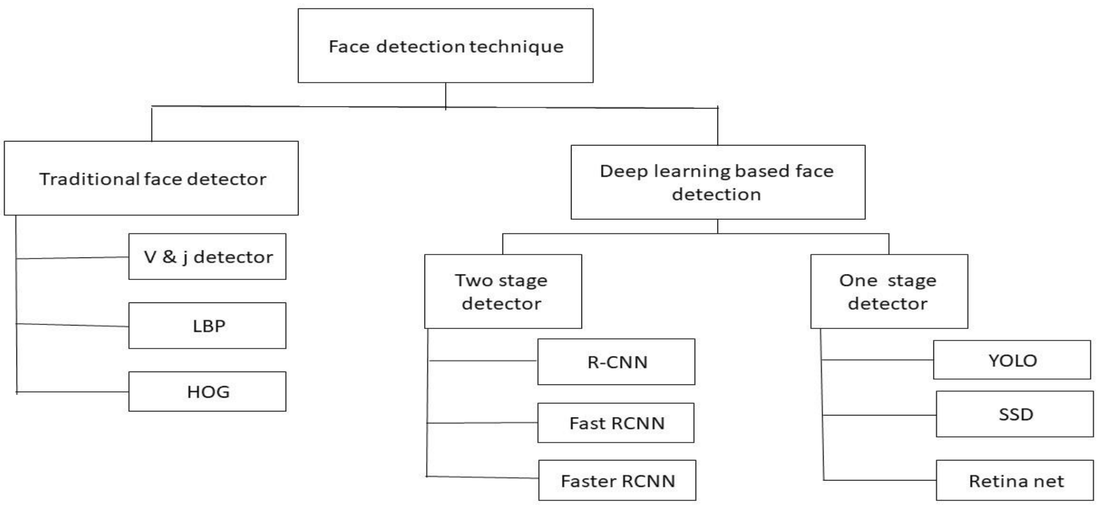
Face detection techniques are essential in computer vision, with a primary focus on identifying and localizing human faces within images and video frames. These techniques are essential for various applications, including photography, security systems, augmented reality, and social media filters. Multiple methodologies exist for detecting faces, including traditional approaches and more modern techniques based on deep learning. Traditional techniques employ feature-based methodologies such as Viola-Jones, which utilizes Hear-like features and cascade

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية
ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

classifiers, and Histogram of Oriented Gradients (HOG), which detects facial regions based on gradients in image intensity. In recent years, face identification has been significantly transformed by deep learning techniques such as Convolutional Neural Networks (CNNs). Models like Single Shot Multi_box Detector (SSD), You Only Look Once (YOLO), and Region-based CNNs (R-CNNs) have demonstrated exceptional precision and efficiency in face detection by acquiring complex patterns and representations directly from the data[7]. Block diagram (1) show samples of face detection techniques.



Block diagram (1) Sample of face detection techniques [21]

4.1Traditional Face Detectors:

The traditional target detection method employs techniques such as Viola-Jones detectors (VJ), LBP, and HOG to identify regions of interest by extracting features from sliding windows. These features are then categorized using classifiers like SVM and Ada Boost. However, manually constructing these features results in higher complexity and limits detection accuracy. The sliding window approach also has drawbacks, including limited detection capability and significant computational demands.[21]

4.1.1 Viola-Jones detectors (VJ)

The Viola-Jones algorithm, developed by Paul Viola and Michael Jones in 2001, revolutionized object recognition by using Haar-like features to quickly identify objects in image. It employs an Adaboost classifier to differentiate between objects and the background, accelerating the detection process. Integral pictures calculate Haar-like characteristics, and the algorithm uses a series of classifiers to focus computation on potential face areas [13].

Figure (3): Haar features an example[22]**.**

4.1.2 Local Binary Pattern (LBP)

It is well-known and referred to by this acronym. The texture descriptor under consideration exhibits diverse applications, including but not limited to environmental modeling, remote evaluation, image recovery, motion analysis, face image analysis, medical image analysis, and face image analysis. The procedure involves partitioning a face picture into distinct regions, wherein Local Binary Patterns (LBP) are concatenated and extracted to form a feature vector. This feature vector is subsequently used as a facial descriptor for further analysis. Ojala and al were early users of the Local Binary Patterns (LBP) algorithm, which uses a 3x3 window to analyze a pixel's neighborhood in an image. The algorithm uses three parameters: central pixel value, grey color, and lower value. The binary code is generated by concatenating 8 1 or 0 values[22].

Let (xc , yc) be pixel positions, 8-bit word resulting in a decimal value that can be expressed as follows:

LBP (xc ,yc) =Σ $S(In - Ic)$ 7 n=0 2n [22]

Where

lc=gray value for central pixel (xc , yc).

ln=gray values of the 8 neighboring pixels.

S (K) ={1 $if\ K \geq 0$, 0 if K<0.

4.1.3 Histogram of Oriented Gradients (HOGs)

It is a feature descriptor that has demonstrated effective use in object and pedestrian identification. It depicts an item as a singular value vector, in contrast to a collection of feature vectors where each vector corresponds to a specific region inside the picture. The computation of HOGs involves employing a sliding window detector that traverses the entirety of an image. The HOG descriptor is calculated for every position, with the picture scale changed to obtain the corresponding HOG feature extraction.[21]

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية
ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

4.2 Deep Learning-Based Face Detection Technique

There are two types of approaches used for target detection based on deep learning, which are as follows:

4.2.1 Two-stage detection models

In This paper aims to provide a comparative analysis of many iterations of the Region-based Convolutional Neural Network (R-CNN) architecture, starting from the original R-CNN to the more advanced Faster_CNN. As shown in the following Block diagram (2):

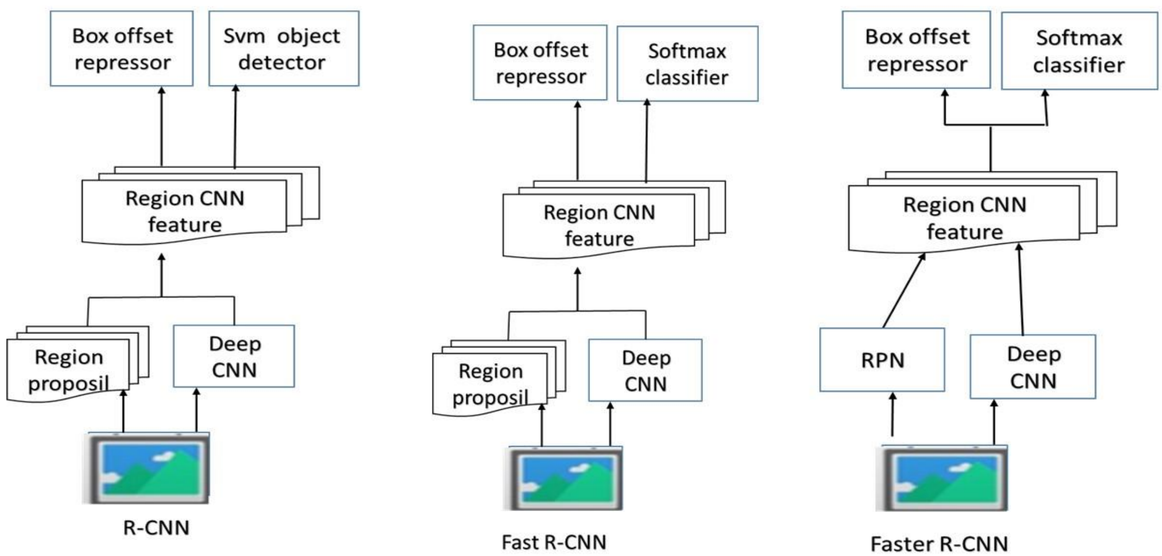
- **Region-based Convolutional Neural Network (RCNN):** The Convolutional Neural Network based on Regions (RCNN) is a two-stage object detection method consisting of three steps: extracting the region, computing CNN features, and classifying the region. The selective search method creates 2,000 clipped and wrapped regions, grouping similar zones based on size, texture, color, and shape. CNN extracts feature from each region, and each class uses a different score and NMS to reject regions with an IOU greater than the learned threshold. Due to its low-level signs, RCNN has drawbacks, such as repeated calculations, time-consuming training and testing, and underperformance in complex image backgrounds **[13,15].**
- **Spatial Pyramid Pooling Networks (SPPNs**): is a deep learning architecture developed to improve the accuracy of traditional convolutional neural networks (CNNs) in handling images of varying sizes. Researchers in 2014 proposed a novel network architecture called SPP-Net, which uses SPP layers to create a fixed-length vector representation compatible with the fully linked convolution layer. This approach uses a singular instance of the entire image for feature extraction, partitioning it into three grids of fixed sizes. The feature vectors are then consolidated and inputted into SVM classifiers and Bounding Box regression. SPP Net outperforms RCNN in precision and computational efficiency but has limitations, such as disregarding preceding layers and multiple stages in training.
- **Fast Region-based Convolutional Neural Network:** Girshic improves RCNN and SPP Net with Fast Region-based Convolutional Neural Network, utilizing whole image input for feature map production using convolution layer[23]

The multi-level pooling layer is eliminated and replaced with a single layer grid. The utilization of the "Region of Interest (ROI)" pooling layer aims to extract a feature vector of fixed length, which is subsequently employed in conjunction with the "Spatial Pyramid layer (SPP)" in cases when only a single pyramid layer is utilized.

The series of "Fully Convolution (FC) layers" processes each feature vector before reaching the output layer. The ultimate output of the FC layers is divided into two sibling layers, namely Softmax and Bounding Box. The softmax layer generates probabilities for the classes of all objects, including one background class for the item. The BBOx layer has produced four real values to create a bounding box around the predicted object. The aforementioned technique does not include distinct training processes for the classifier and bounding box regression. The proposed method effectively leverages the benefits of both RCNN and SPP-Net. However, its computational speed is slightly hindered due to the inclusion of proposal detection. However, this practice results in cost savings by reducing additional expenses for storage space, while also enhancing accuracy and efficiency.

- **Faster Region-based Convolutional Network (Faster R-CNN)**:

The approach is a computer vision technique that aims to improve the speed and accuracy of object detection in images. Various techniques are employed to generate the candidate boxes, such as selective search and edge box. Nevertheless, the efficacy of the object detection strategy still needs to be improved by these strategies. In 2015, a novel approach called Region Proposal Network (RPN) was presented by S. Ren, K. He, and R. Girshick et al[3]. in their work on Faster RCNN. This method aimed to generate regions, as documented in reference. The RPN algorithm employs the sliding window technique on the feature map to produce the bounding box for each object and the corresponding bounding box score.



Block diagram (2):Comparison of different version of R-CNN from RCNN to Faster-RCNN [24]**.**

4.2.2 One-Stage Detection Models

Unlike the two-stage target identification approach, the one-stage target detection method consistently performs dense sampling in various picture positions instead of carrying out regional suggestions. In sampling, various aspect ratios and scales are employed with the regression technique Boundary box and classification probability to extract the target from the CNN feature map directly.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

- **You Only Look Once** The YOLO (You Only Look Once) algorithm, introduced by Redmon, Divvala, Girshick et al. in 2015[4], uses a fixed grid detector for object recognition and verification. It uses a unified neural network to detect objects across an image, partitioning it into predetermined sections for computation of probability and bounding boxes. A singular Convolutional Neural Network estimates the probability of different classes across bounding boxes.

The training process uses complete photos, and the algorithm uses convolutional layers with a grid system for predictions.

several iterations of YOLO include YOLOv2, YOLOv2 tiny network with ear detection , YOLOv3 ,YOLOv4, and YOLOv5 [21]**.**

- **Single Shot Multi Box Detector:** In 2015, Liu, Anguelov, Erhan, and colleagues proposed a new approach known as SSD (Single Shot Multi Box Detector) [25] .The Single Shot Multi Box Detector (SSD) is a deep learning method that uses a multi-reference and multi-scale representation approach to enhance the precision of detecting tiny objects.

Unlike previous methods, SSD operates on the uppermost layers and has high accuracy in real-time fire detection[26]. The SSD architecture consists of a backbone model and an SSD head, which are combined to produce a bounding box (BB) that encompasses the detected object. This highly efficient object detection model can rapidly detect many categories.

- **Retina net:** Lin, Goyal, Girshick, and colleagues propose a new loss function as an alternative to the conventional cross-entropy loss in this method. This new loss function addresses the class imbalance challenge during the model's training[27]**.** RetineNet has developed a single-stage object detection model with efficient computational speed and high precision for detecting densely distributed and smaller-scale objects. The system comprises a

backbone architecture and two subnetworks, which perform convolutional operations, object categorization, and bounding box regression. This approach enhances the efficiency of a one-stage detector, as it accommodates input images of any size and performs convolutional operations.

6.THE DATASETS USED FOR FACE DETECTION

This section provides a concise explanation of the significance of using datasets for face detection research. This text emphasizes the significance of datasets in the training and evaluation of face identification algorithms and how they contribute to the progress of the field. The commonly used face detection datasets in academic research. It encompasses popular datasets such as:

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

- **WIDER FACE Dataset**: The WIDER FACE database (Yang et al., 2016), serves as a standard benchmark for face detection. The dataset has a total of 32,203 photos, within which t,here are 393,703 faces displaying significant variations in terms of scale, position, and occlusion. The structure of this database is structured according to 61 distinct event classes[28].
- **The Face Detection Data Set and Benchmark (FDDB):** Jain & Learned-Miller [33] was created to study the challenge of identifying faces in unconstrained conditions. This involves addressing significant changes in appearance caused by factors like variations in position, low image quality, differences in face size, and variations in brightness. The database consists of annotations for a total of 5,171 facial pictures collected from a collection of 2,845 photographs obtained from the Faces in the Wild dataset. The database was employed to assess the efficacy of the developed model for face detection. [29].
- **Pascal Faces dataset**: The Pascal Faces dataset is comprised of 1335 labelled faces on a total of 851 images that were retrieved from the Pascal VOC dataset [8].
- **AFW dataset:** The AFW dataset, developed by Z. Zhang et al. in 2020, is a special dataset created to assess the performance of face detection algorithms under challenging and uncontrolled circumstances. The dataset comprises photos exhibiting substantial disparities in terms of stance, emotion, and illumination. The AFW dataset has 473 labeled faces distributed among 205 different images. [23].

7. CONCLUSIONS

In recent times, there has been a significant focus from researchers in the field of computer vision on the topic of face detection. This survey has comprehensively examined different methodologies for detecting faces, emphasizing detection techniques' difficulties, uses, and progression from conventional to deep learning-based methods. Additionally, it offers valuable information about frequently utilized datasets for facial detection, providing a thorough analysis of their distinctive features. The study is organized sequentially, beginning with an introduction to evaluating noteworthy prior research, and the difficulties encountered by face detection systems. This is followed by exploring various applications, outlining detection approaches, and analyzing commonly used datasets.

The progression of face identification technology from conventional classifiers to deep learning models signifies a substantial advancement in tackling the intricacies of face detection in many settings. Despite existing challenges, the ongoing progress in deep learning presents hopeful solutions that improve the precision and effectiveness of face identification systems in various applications.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com
مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Conflict of interests.

There are non-conflicts of interest.

References

[1] K. A. Majeed, Z. Abbas, M. Bakhtyar, J. Baber, I. Ullah, and A. Ahmed, “Face Detectors Evaluation to Select the Fastest among DLIB, HAAR Cascade, and MTCNN,” *Pakistan J. Emerg. Sci. Technol.*, vol. 2, no. 1, pp. 50–62, 2021.

[2] L. Alzubaidi *et al.*, “Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions,” *J. big Data*, vol. 8, pp. 1–74, 2021.

[3] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” *Adv. Neural Inf. Process. Syst.*, vol. 28, 2015.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[5] W. Liu *et al.*, “Ssd: Single shot multibox detector,” in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, Springer, 2016, pp. 21–37.

[6] W. Yang, L. Zhou, T. Li, and H. Wang, “A Face Detection Method Based on Cascade Convolutional Neural Network,” *Multimed. Tools Appl.*, vol. 78, no. 17, pp. 24373–24390, 2019, doi: 10.1007/s11042-018-6995-0.

[7] N. Zhang, J. Luo, and W. Gao, “Research on face detection technology based on MTCNN,” *Proc. - 2020 Int. Conf. Comput. Network, Electron. Autom. ICCNEA 2020*, pp. 154–158, 2020, doi: 10.1109/ICCNEA50255.2020.00040.

[8] Z. Zhang, W. Shen, S. Qiao, Y. Wang, B. Wang, and A. Yuille, “Robust face detection via learning small faces on hard images,” *Proc. - 2020 IEEE Winter Conf. Appl. Comput. Vision, WACV 2020*, pp. 1350–1359, 2020, doi: 10.1109/WACV45572.2020.9093445.

[9] H. Mliki, S. Dammak, and E. Fendri, “An improved multi-scale face detection using convolutional neural network,” *Signal, Image Video Process.*, vol. 14, pp. 1345–1353, 2020.

[10] S. Hou, D. Fang, Y. Pan, Y. Li, and G. Yin, “Hybrid Pyramid Convolutional Network for Multiscale Face Detection,” *Comput. Intell. Neurosci.*, vol. 2021, 2021, doi: 10.1155/2021/9963322.

[11] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, “YOLO-face: a real-time face detector,” *Vis. Comput.*, vol. 37, no. 4, pp. 805–813, 2021, doi: 10.1007/s00371-020-01831-7.

[12] Q. Xu, Z. Zhu, H. Ge, Z. Zhang, and X. Zang, “Effective face detector based on YOLOv5 and superresolution reconstruction,” *Comput. Math. Methods Med.*, vol. 2021, pp. 1–9, 2021.

[13] Y. Liu, R. Liu, S. Wang, D. Yan, B. Peng, and T. Zhang, “Video Face Detection Based on Improved SSD Model and Target Tracking Algorithm,” *J. Web Eng.*, vol. 21, no. 2, pp. 545–567, 2021, doi: 10.13052/jwe1540-9589.21218.

[14] N. Kulkarni, D. Mantri, P. Pawar, M. Deshmukh, and N. Prasad, “Occlusion and Spoof Attack Detection using Haar Cascade Classifier and Local Binary Pattern for Human Face Detection for ATM,” *AIP Conf. Proc.*, vol. 2494, no. October, 2022, doi: 10.1063/5.0107262.

[15] T. H. Obaida, A. S. Jamil, and N. F. Hassan, “Real-time face detection in digital video-based on Viola-Jones supported by convolutional neural networks,” *Int. J. Electr. Comput. Eng.*, vol. 12, no. 3, pp. 3083–3091, 2022, doi: 10.11591/ijece.v12i3.pp3083-3091.

[16] D. Mamieva, A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, “Improved face detection method via learning small faces on hard images based on a deep learning approach,” *Sensors*,

vol. 23, no. 1, p. 502, 2023.

[17] P. N. Achyutha, S. Hebbale, and V. Vani, "Real time COVID-19 facemask detection using deep learning," *Int. J. Health Sci. (Qassim).*, vol. 6, no. March, pp. 1446–1462, 2022, doi: 10.53730/ijhs.v6ns4.6231.

[18] H. Fei, B. Tu, Q. Chen, D. He, C. Zhou, and Y. Peng, "An overview of face-related technologies," *J. Vis. Commun. Image Represent.*, vol. 56, no. September, pp. 139–143, 2018, doi: 10.1016/j.jvcir.2018.09.012.

[19] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," *Artif. Intell. Rev.*, vol. 52, no. 2, pp. 927–948, 2019, doi: 10.1007/s10462-018-9650-2.

[20] A. N. Razzaq, R. Ghazali, N. K. El Abbadi, and M. Dosh, "A Comprehensive Survey on Face Detection Techniques," *Webology*, vol. 19, no. 1, pp. 613–628, 2022, doi: 10.14704/web/v19i1/web19044.

[21] J. Kaur and W. Singh, "Tools, techniques, datasets and application areas for object detection in an image: a review," *Multimed. Tools Appl.*, vol. 81, no. 27, pp. 38297–38351, 2022, doi: 10.1007/s11042-022-13153-y.

[22] B. T. Devi and S. Shitharth, "Multiple Face Detection Using Haar - AdaBoosting, LBP-AdaBoosting and Neural Networks," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1042, no. 1, p. 012017, 2021, doi: 10.1088/1757-899x/1042/1/012017.

[23] R. Girshick, "Fast R-CNN," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, pp. 1440–1448, 2015, doi: 10.1109/ICCV.2015.169.

[24] A. Ayanzadeh, "A Study Review : Semantic segmentation with Deep Neural Networks," no. March, 2019, doi: 10.13140/RG.2.2.12534.04160/1.

[25] W. Liu *et al.*, "Ssd: Single shot multibox detector. in European conference on computer vision. Springer," 2016.

[26] S. Wu and L. Zhang, "Using popular object detection methods for real time forest fire detection," in *2018 11th International symposium on computational intelligence and design (ISCID)*, IEEE, 2018, pp. 280–284.

[27] A. N. Azhar and M. L. Khodra, "Fine-tuning Pretrained Multilingual BERT Model for Indonesian Aspect-based Sentiment Analysis," *2020 7th Int. Conf. Adv. Informatics Concepts, Theory Appl. ICAICTA 2020*, pp. 2980–2988, 2020, doi: 10.1109/ICAICTA49861.2020.9428882.

[28] S. Yang, P. Luo, C.-C. Loy, and X. Tang, "Wider face: A face detection benchmark," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5525–5533.

[29] H. Mliki, S. Dammak, and E. Fendri, "An improved multi-scale face detection using convolutional neural network," *Signal, Image Video Process.*, 2020, doi: 10.1007/s11760-020-01680-w.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية
ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

JOURNAL OF UNIVERSITY OF BABYLON
For Pure and Applied Sciences (JUBPAS)
Vol.32; No.2.| 2024

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com ISSN: 2312-8135 | Print ISSN: 1992-0652

خلاصة

شهد مجال اكتشاف الوجه تقدمًا كبيرًا، يُعزى في الغالب إلى التقدم في التعلم الآلي ومنهجيات التعلم العميق. وقد أدت هذه التحسينات إلى تغيير جذري في هذا المجال. تُستخدم خوارزميات اكتشاف الوجه على نطاق واسع في العديد من المجالات، مثل تطبيقات الأمان والمراقبة والشبكات الاجتماعية. لقد بحث هذا الاستطلاع بشكل شامل في منهجيات مختلفة للتعرف على الوجه، مع التركيز على التحديات والتطبيقات والتطور من الأساليب التقليدية إلى الأساليب القائمة على التعلم العميق. علاوة على ذلك، فإنه يوفر رؤى مهمة حول مجموعات البيانات شائعة الاستخدام للكشف عن الوجه، بما في ذلك الفحص الشامل لخصائصها الفريدة. تم تنظيم الدراسة بشكل تسلسلي، بدءًا بمقدمة لتقييم الأبحاث السابقة المهمة والتحديات التي تواجهها خوارزميات اكتشاف الوجه. بعد ذلك، يتعمق البحث في الاستخدامات العملية المختلفة، ويحدد طرق تحديد وتدقيق مجموعات البيانات المستخدمة بشكل متكرر. يمثل الانتقال من المصنفات التقليدية إلى نماذج التعلم العميق في تقنية التعرف على الوجوه تحسنًا كبيرًا في معالجة تعقيدات اكتشاف الوجوه في البيئات المختلفة. على الرغم من العقبات الموجودة الآن، فإن التقدم المستمر في التعلم العميق يقدم حلولاً واعدة تعزز دقة وكفاءة أنظمة التعرف على الوجه في مجموعة واسعة من التطبيقات.

الكلمات المفتاحية: كشف الوجه، التعلم الآلي، التعلم العميق، المجال المكاني، CNN.