Journal of Applied Engineering and Technological Science
Vol 5(2) 2024: 813-829

THE ROLE OF ARTIFICIAL INTELLIGENCE IN PROVIDING PEOPLE WITH PRIVACY: SURVEY

Salar Adil Raees[1], Mohammed S.H.Al-Tamimi[2*]
Computer Science Department, College of Science, University of Baghdad, Iraq[12]
salar.raees2101m@sc.uobaghdad.edu.iq[1], m_altamimi75@yahoo.com[2*]

Received : 30 December 2023, Revised: 07 February 2024, Accepted : 16 March 2024
Corresponding Author

ABSTRACT
Images privacy involves assessing the amount of information leakage from images, assessing risks associated with identification, and examining controls on this information. It was discussed various types of protection available and most commonly used in providing privacy to a person in images, including single-stage and two-stage detection algorithms. The results of each algorithm are organized in detailed tables, and the [YOLO] algorithm expands on all versions. The paper also clarifies the dataset used for testing the algorithms and its relevance to achieving desired results. It presents a comprehensive understanding of the process of detecting persons in digital images and assesses various tools and algorithms for recognizing persons, faces, and identities. It added an extensive examination of the several methods used to identify persons in digital images, with a specific emphasis on safeguarding their privacy. The task at hand is assessing various face recognition and identification tools and algorithms, with a specific emphasis on those that exhibit superior accuracy and efficiency in presenting outcomes. The study concluded that using the yolov8 algorithm in conjunction with blurring techniques effectively conceals individuals' information in digital images while maintaining the integrity of the overall image. The research paper's implications and information can practically contribute to the development of algorithms for detecting and protecting people in digital images, as well as the development of applications in this field. Theoretically, it can enhance understanding of the process of detecting and protecting people, and potentially contribute to the development of new theories in the field of protection and discovery.
Keywords : *Privacy, Face, Person, Methods, Dataset, AI.*

1. Introduction

Facial recognition technology is revolutionizing the analysis of images, enhancing body privacy but also raising concerns about potential misuse. One issue is the infringement of individuals' privacy through the unauthorized collection and utilization of their data from digital images, as well as the sharing of this data with third parties without consent. Another concern is the discriminatory application of facial recognition technologies to differentiate between specific groups of people. Personal identification technologies have the potential to be used for surveillance purposes, enabling the monitoring of individuals and encroaching upon their privacy and freedom of movement. Regarding security, personal information might be used for activities such as impersonation, fraudulent acts, or the dissemination of deceptive information. Previous research in the field of digital images used object detection (Anushka et al., 2021), person detection (W. Wang, 2020), and medical face mask detection (Gonzalez Dondo et al., 2021). this is all in the field of detection, but the field of protection includes Face Detection and protection and body detection and protection (Wen et al., 2022) (Shifa et al., n.d.) . These technologies aim to hide individuals' identities or conceal sensitive information while preserving the public benefit of images for various purposes. Failure to protect people's data can have ethical consequences, such as violating privacy, resulting in loss of control over personal information and exposure to surveillance. It can also lead to psychological harm, fear of being tracked, loss of sense of security, reputational damage, and legal liability. that State institutions and companies are legally responsible for safeguarding individuals' data. And Certain parts of the globe, including certain nations in the Middle East, lack legislative requirements for safeguarding personal information, unlike other countries such as those in Europe (GDPR), America (CSPA), Canada (PIPEDA), Australia (Privacy Act), and Japan. The goal of protection can be achieved by various methods, such as using different types of blurring to achieve the desired effect without compromising the image's features (Pulfer, 2019). Other methods include (masking, pixelation, encryption, and scrambling) (Zhong & Deng, 2023) (Lander et al., 2001) (Aribilola et al., 2022) (Hosny et al., 2022). However, one of these methods' weaknesses is that it may distort the image's environment or draw attention because it conceals something significant from the viewer (Rakhmawati et al., 2018) . And The study of protecting individuals in images is subject to several limitations. These include the challenge of acquiring adequate data for a comprehensive analysis due to the limited availability of such data and the privacy concerns surrounding the subjects. Additionally, the high costs associated with skin scanning and ethical restrictions, such as a possible encroaching of privacy, further hinder the study. Privacy protection aims to keep information that a person wishes to remain private out of the public eye. Call it visual privacy protection while sharing images and videos.

Additionally, which delicate information or area of interest has to be secured is a crucial question regarding visual privacy. Even if the face is often the only thing concealed in these fragments, visual privacy is nevertheless violated. Even if a person's face is hidden, there may still be other components in the image that allow for personal identification, such as using prior information and deduction channels (Padilla-López et al., 2015) (Intelligence and Neuroscience, 2023).

Digital images of faces, users can share their images via various channels and social networking sites. Globally, users of the biggest social networking sites share large numbers of face images, according to statistics. These computerized images typically hold a plethora of private, delicate data. Furthermore, due to the compromise of personal privacy, incalculable costs will result from the collection and analysis of the data by attackers (Chettri & Borah, 2015). Image data privacy protection frequently depends on approaches like k-anonymity, Access management, and privacy cryptography. It proposed the method based on the anonymization mechanism (Xiao & Tao, 2006) (Shetty & Jogi, 2019). Preserving the confidentiality of individuals depicted in images used to identify persons numerous advantages, such as fostering trust among users, who will feel secure when utilizing applications that provide this form of safeguarding. Additionally, it promotes transparency in the utilization and dissemination of data. Safeguarding the privacy of digital images also stimulates the creation of novel image-centric applications that prioritize user privacy. Furthermore, this feature empowers users to regulate their personal images and safeguard this data from unauthorized use.

This survey will discuss face and body protection in images so that personal information about the subject is protected. The process of face and body protection involves two main steps: the first is to identify the area that needs to be guarded, and the second is to hide information about the area that needs to be secured to meet the necessary protection standards. Also, note that most research only protects the face and not the entire body.

2. Literature Review

Numerous research have been conducted on this issue. The issue of individuals' privacy and identity remains a prominent concern that requires resolution. This paper proposes a solution to the problem of safeguarding individuals' privacy by using an efficient technology that encrypts specific images. The encryption process employs a chaotic approach to transferring and storing these images, with a particular focus on selectively encrypting the face area. This technology enhances the efficiency of encryption and decryption operations, reducing computational requirements and significantly improving overall performance. The foundation of this algorithm's operation is the Viola-Jones detection concept. The process involves extracting the facial features from the picture and encrypting them using a permutation-diffusion architecture (Kouadra et al., 2023) . This research presents a method for safeguarding the privacy of persons in videos by using block scrambling-based encryption and DCNN-based face identification. The yolov3 method is used to do Face Detection in images. and applying quick block scrambling to the identified faces. Next, encryption is applied, which obfuscates the data using a unique key that is produced in a random and unpredictable way. The findings demonstrate that the suggested technology effectively identifies and safeguards individuals' facial features without any data breaches and is resilient against potential threats (Hosny et al., 2022). Autonomous pedestrian identification is crucial for a range of computer vision applications such as intelligent video surveillance, advanced traffic monitoring, and obstacle detection in smart urban environments. The research suggests including an anti-residual module into the robust Enhanced YOLOv3+ network to enhance feature extraction. The network is optimized by minimizing bounding box loss error and trained on the Pascal VOC-2007+12 dataset using just extracted pedestrian pictures.

The INRIA Pedestrian dataset is used with the PASCAL VOC-2007+12 dataset. The dataset contains a training set of 614 positive images with 1237 pedestrians and a testing set of 228 positive images with 589 pedestrians. The images in the INRIA Pedestrian dataset are all 64 x 128 pixels in resolution. The INRIA Pedestrian dataset, like the Pascal VOC-2007+12 Pedestrian dataset, has images with intricate backdrops, varying lighting conditions, variable levels of obstruction, diverse human poses, and individuals in various outfits (Murthy & Hashmi, 2020). The research suggests an enhanced pedestrian recognition technique using Faster RCNN to increase the detection accuracy of small-scale pedestrians in images. The technique utilizes a deep convolutional neural network (CNN) to extract visual data and then selects suitable pedestrian zones by clustering and a regional Recommendation Network (RPN). A feature fusion technique using cascade is suggested to boost the semantic information of the network via the integration of high-level and low-level characteristics.

The Online Hard Example Mining (OHEM) approach is used to train data with significant loss in order to enhance detection performance by mitigating the imbalance between positive and negative samples. The approach enhances the Faster RCNN framework by refining the RPN region network via the use of the k-means clustering algorithm to minimize computational load. Feature fusion and the OHEM algorithm are used to improve pedestrian detection performance. The proposed approach performs better than Faster RCNN in recognizing tiny target pedestrians, as shown by experimental results on PASCAL VOC2007 and INRIA pedestrian datasets (Shao et al., 2021) . The research explores the use of the Yolov5 model for object identification in many scenarios, including flower picture classification, face mask recognition, breast tumor diagnosis, and face detection. The Yolov5 model is evaluated against other models such as Faster R-CNN, R-FC, and SSD, using precision as the performance measure. Non-Max Suppression is a method used in object identification to choose the most optimal bounding box. The process includes comparing objectiveness ratings and eliminating bounding boxes with IOUs above 50. The report presents experimental data and performance indicators, including accuracy, recall, and mAP, to assess the effectiveness of the Yolov5 model across several applications. The report ends with suggestions for further research in the area of object detection with the Yolov5 model (Tahir et al., 2023). The research focuses on real-time pedestrian detection using tiny mobile devices with limited processing resources.

The study introduces three techniques: Enhanced Local Binary Pattern (LBP) characteristics and Adaboost classifier. Enhanced Histogram of Oriented Gradients (HOG) features using a Support Vector Machine (SVM) classifier. The study discusses the difficulties associated with recognizing pedestrians, including the wide range of appearances due to factors like stance, clothes, occlusions, weather conditions, and background distractions. The LBP features and Adaboost classifier was selected for their capability to extract contour features efficiently, while the HOG features and SVM classifier, although popular, are slower because of their extensive search process. The research enhances the efficiency of the exhaustive search in HOG and SVM to increase speed (Weng, 2023).

3. Type of Visual Privacy Protection

There are standard practices for protecting personal data privacy in pictures and videos. The six standard editing techniques are masking, encrypting, pixelation, black-box, blurring, and scrambling. Blurring is the consequence of working with a Gaussian formula on an image. The blurring process includes using surrounding pixels to help change the pixel value of the picture, depending on the function being used. Thus, blurring may be used to conceal the area the site is used to hide faces and license plates using the black box approach with Google Street View Apps. Faces and other important information are recognized and localized first, and then the area is replaced with black (or white) rectangles, ellipses, or circles(Shifa et al., n.d.) (Rakhmawati et al., 2018).

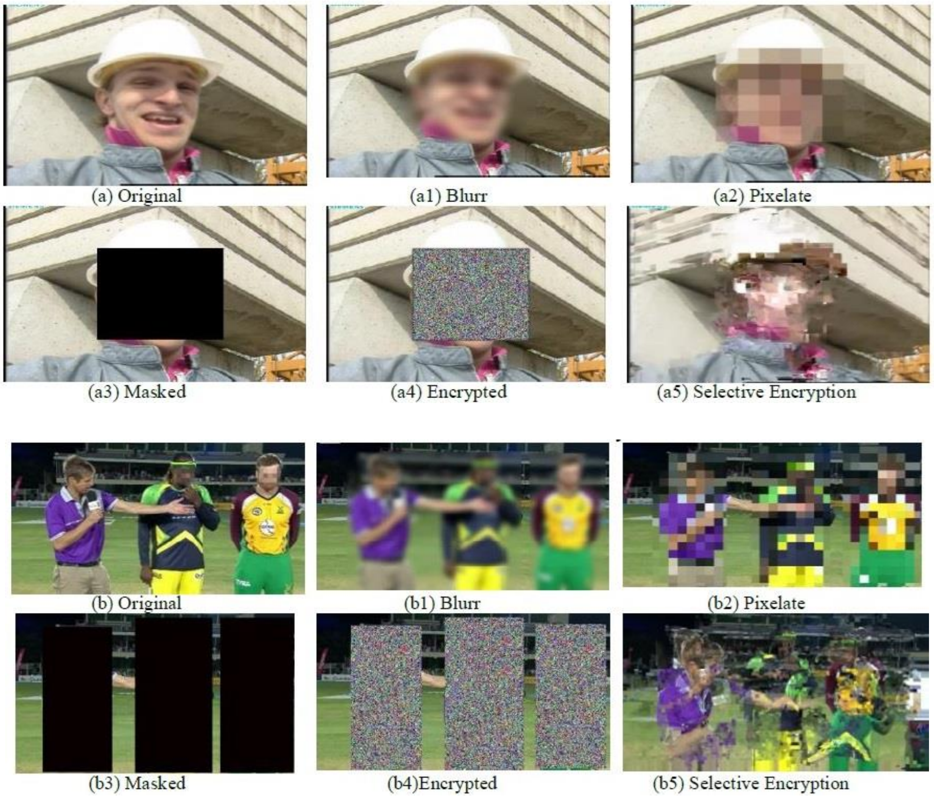

Fig. 1. Images that illustrate privacy(Shifa et al., n.d.).

The pixelation method involves taking the mean of a specific size's pixel number. Moreover, employs an 8 × 8-pixel size. Every pixel that is currently there will be substituted using the mean pixel value in its entirety. Thus, the resolution of sensitive data may be reduced to the point where it becomes unreadable. This technique is often used for censorship of news and documentary television programs. This is implemented to protect victims, eyewitnesses, suspects, and other significant elements of proof before they are processed further. When pixelation is absent, it is possible to reconstruct the original picture and utilize non-pixelated pixels to adjust any differences between big pixels (Adarsh et al., 2020). Including reduced pixel values into still images is not difficult in this case. Continuous or random colour is employed wherever feasible to maintain the ability to exhibit outstanding visual qualities. Other masking techniques include determining the region of interest (ROI) and then using a solid colour to conceal sensitive parts. The most commonly used privacy protection method is data encryption. The encryption key is used to localize sensitive data; if the receiver lacks the necessary decryption key, the sensitive data will not be visible to it (Redmon et al., 2016). Put differently, encrypting data may protect an individual's identity and actions. We may work with operators and use a few decryption keys to safeguard the data. Scrambling is the last method in this area, which, depending on the kind of material, may be implemented via code stream domains, spatial domains, or transformation domains (Shifa et al., n.d.) (Rakhmawati et al., 2018).

3.1 Masking

This technology offers specific mechanisms for concealing information displayed in digital images, allowing for meticulous control. However, it requires manual intervention for implementation and may not be effective when dealing with dynamic objects in video presentations. One notable advantage is the ability to exercise precise control and explore creative options, although it is a more complex technique compared to other alternatives.

3.2 Encrypting

The encryption technology is dependent on the utilization of a key that plays a crucial role in facilitating the encryption of images, rendering them unreadable without decryption. This technology offers robust protection, albeit with a level of complexity in its usage, as the key is required for decrypting the encrypted images in order to make them legible. The encryption procedure is not feasible for daily uses. It is characterized by offering security measures to safeguard and restrict unwanted access to sensitive information in images. However, one of its drawbacks is its complexity in use. Additionally, decryption of the code always necessitates a key, which may be compromised by malicious entities. Encryption is appropriate for securing sensitive images, particularly those of a medical nature. Nevertheless, her visa is extensive and her privacy is very intricate.

3.3 Pixelation

This technique works by increasing the dimensions of the chosen pixels inside the image. As a result, it introduces a specific sort of noise that is known for its straightforward operation, reduced level of detail, and limited effectiveness when dealing with tiny objects. One advantage of this technology is its simplicity and ability to produce a specific type of noise. However, it is not effective for small objects and can make the body appear bulky. It is suitable for disguising public locations and movement patterns, and it has a strong impact on images with moderate privacy and low complexity.

3.4 Black-Boxing

This method distinguishes itself from previous ways by using a black layer to obscure the apparent parameters of digital images, thereby concealing the information. However, this approach results in a conspicuous alteration of the visual surroundings, drawing attention to the concealed region. This technology is user-friendly, however, it also piques interest and curiosity about the concealed features in digital images. It is well-suited for videos and apps that need censoring since it has a significant effect and is not too difficult.

3.5 Blurring

The primary objective of this technique is to regulate the degree of noise present in the digital image. This feature is advantageous for concealing information shown in digital images from an artistic perspective. This feature is known for its capacity to provide freedom in selecting the desired degree of noise, while also effectively preserving the integrity of the visual environment and avoiding drawing attention to areas where noise may be present. One of his limitations is that he does not fully conceal himself in the images. However, the impact of the distortion varies depending on the level of intensity used, exhibiting a moderate level of specificity and a medium level of complexity.

3.6 Scrambling

This technology has a profound impact on the content of the image or the specific area that has been edited, making it extremely challenging to retrieve information. It is distinguished by its capability to disable the content of the digital image and offers excellent privacy. However, one drawback is its high cost and the difficulty of reversing the changes without the use of a key. It is ideal for entirely concealing very sensitive information that has a significant impact and is highly intricate.

4. The evolution of privacy protection technologies

Preserving the privacy of individuals in digital images is a crucial technology that aims to safeguard personal information. It is imperative to continuously develop this technology, particularly with the advancements in machine learning. Neural networks play a significant role in achieving these objectives by training them to conceal facial identities, license plates, or any other specific information in an image. These networks have the ability to substitute crucial regions with other options and retain confidentiality. Privacy in digital images may be achieved by introducing a kind of noise specifically targeted at the crucial areas of the images. These technologies have the ability to alter the image in a manner that is undetectable by humans and cannot be recognized by face recognition systems. In contrast, machine learning algorithms has the capability to autonomously identify certain items inside digital images. This facilitates the process of automated retouching, resulting in time savings when compared to manual masking. The use of these sophisticated technologies plays a significant role in the development of viable solutions in the realm of visual information.

5. Applications of privacy

Personal information protection technologies are commonly employed to safeguard sensitive data. This can be achieved through program and news censorship, as well as through social media platforms like Facebook and other applications. These technologies involve obscuring crucial information in videos and digital images. Additionally, they can be utilized in the healthcare sector to block information that could potentially reveal the identity of patients in medical images. This is particularly important when sharing these images among doctors, and it is crucial to ensure a high level of protection. The significance of this technology in the legal domain cannot be underestimated, as it ensures the confidentiality of witness identities and the preservation of their safety while identifying them in criminal evidence.

6. Obstacles and Constraints

The preservation of visual privacy is consistently confronted with significant and ongoing obstacles due to the continuous advancement of technology, which diminishes its efficacy. These challenges encompass the difficulty in detecting accurate information for individuals. However, this issue can be resolved by integrating multiple technologies, such as Face Detection, along with analysis techniques, to yield effective results and solutions. Alternatively, use deep learning algorithms that may be specifically designed to achieve the intended outcomes. It is feasible to strike a balance between preserving privacy and facilitating the usage of images by doing research that promotes development and offers choices to consumers.

7. User Perception and Acceptance

The user's approval and awareness of the need to safeguard privacy in images and videos is a very delicate and crucial matter in order to guarantee their creation. Understanding the functioning of data security measures and how to safeguard it, as well as being knowledgeable about the implementation specifics of various protective technologies, is crucial. Regarding user choices, they are related to the level of privacy desired. Some users may want to publish their movies or images without any restrictions, while others may opt to apply protective measures to their external postings. These are the choices that users must have when sharing anything. It is important to consider the cultural norms surrounding the content being shared, such as images or videos, and how different cultures may influence these norms. Some cultures may have restrictions on sharing certain types of digital images or videos, while others may be more open to sharing this kind of information. Ultimately, it is essential for all stakeholders, including developers, manufacturers, and consumers, to collaborate in order to ensure that these technologies are user-friendly, efficient, and aligned with cultural norms.

8. Methods

In order to safeguard people's privacy in images and videos effectively, object identification algorithms should be used with face and body privacy strategies. Combining these technologies, object detection can be done simultaneously as face and body anonymization or blurring to avoid identification. This method can benefit security, surveillance, and medical imaging (Ning et al., 2017) (Shetty et al., 2021).

Object detection with anonymization: Using this method, items in the picture or video are first identified and located using object detection algorithms (Dhillon & Verma, 2020). Subsequently, face and body privacy measures, such as blurring or anonymizing, are used to identify identifiable faces and bodies (Khaled & Al-Tamimi, 2021). This method maintains individual privacy protection while enabling item identification.

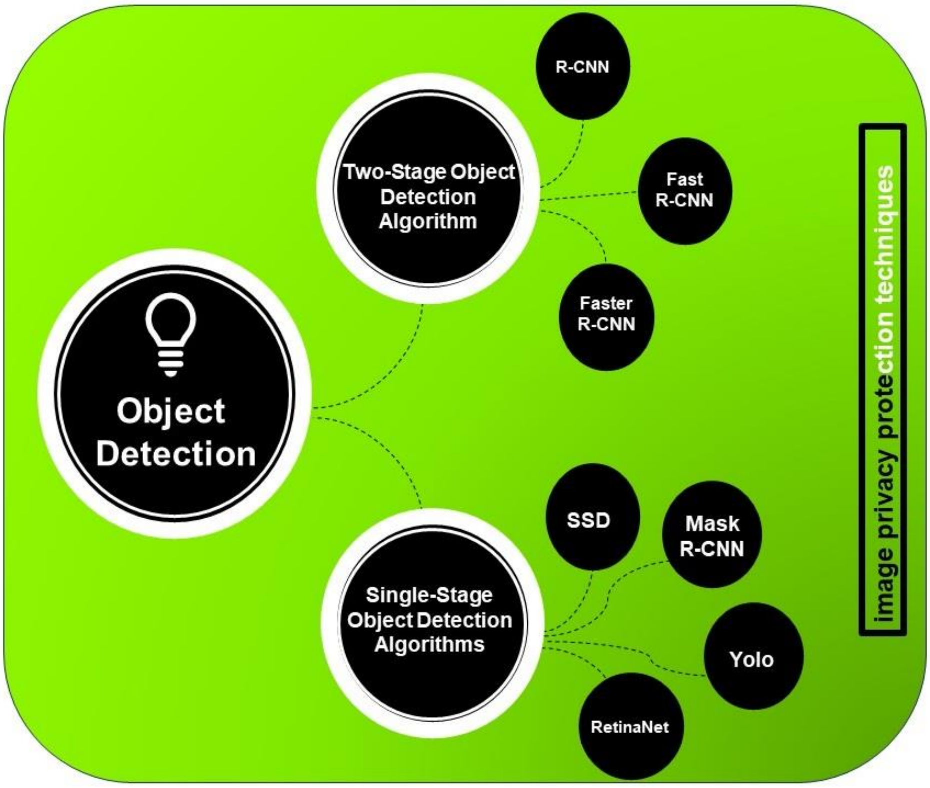

Fig. 2. A classification of image privacy protection techniques (Rakhmawati et al., 2018).

8.1 Single-Stage Object Detection Algorithms

Single-stage detectors only one pass through the neural network, predict all of the bounding boxes simultaneously, and use neural networks of the feedforward type (Mittal et al., 2020) (Patel & Salot, 2021) Because there is a decrease in computing cost and an increase in performance, it is substantially faster. We may transform the two-stage model into a single-stage model by making various adjustments and removing the model's pipelines, which allows the model to operate at a high processing speed (Abed et al., 2023) (Rahma & Salman, 2022). YOLO, SSD, and Retinanet are the most popular types of one-stage object detectors.

8.1.1 The Single-Shot Detector (SSD) model

Was released in 2015. SSD is a single-stage object detection methodology that uses a single pipeline to identify objects. In order to increase precision, The SSD model is adaptable and can be added to systems that need object detectors (Girshick, 2015). The SSD's higher resolution layers are in charge of identifying small things, but they also have some unimportant properties that make them less helpful and useless for object detection (Amer et al., 2022). SSD's single-stage architecture allows for speedier performance and end-to-end training of the entire model. Compared to YOLOv1 and other two-stage models, SSD makes more predictions and can accommodate different aspect ratios. When this model is compared to the two-staged models, its accuracy is slightly lower (Bai, 2020); Table 1 summarizes the ssd algorithm.

Table 1 - SSD algorthem summarize.

Type	advantage	Disadvantages	Domain
SSD	• Fast detection speed • High accuracy • Scalability • Simplicity	• Lower accuracy than some other models • Less robust to noise • It can be less accurate for small objects:	• Image and video surveillance • Autonomous vehicles • Image retrieval • Product recommendation • Medical imaging

8.1.2 YOLO (You Only Look Once):

YOLO is a method that unifies different parts of recognizable elements into a single neural network. It predicts each bounding box by using the features present in the original image. Each image is split into an S×S grid, and each grid predicts bounding boxes along with A degree of trust for each box (He et al., 2020). Whether or not the correct image is enclosed by the box that borders it, the confidence shows its accuracy. YOLO has the advantage of being extremely quick and able to be used in a real-time setting, but it only finishes one kind of class in one grid (Abood, 2023) ; Table 2 summarizes the YOLO algorithm.

Table 2 - YOLO algorithm summarization.

Type	advantage	Disadvantages	Domain
YOLO	• Speed • Accuracy • Simplicity	• Accuracy vs. speed trade-off • Localization errors • Small object detection:	• Video Surveillance • Self-Driving Cars • Augmented Reality • Robotics and Automation • Drone Surveillance and Inspection • Medical Imaging Analysis • Sports Analytics and Performance Tracking • Traffic Monitoring and Management • Retail and Product Analysis • Wildlife Monitoring and Conservation

8.1.3 RetinaNet

Retin a Net is a single, cohesive network comprising two task-specific subnetworks plus a backbone network. As an off-the-shelf convolutional network, the backbone performs a convolutional feature map computation for the whole input image. Using the output of the backbone, the first subnet carries out convolutional object classification, while the second subnet carries out convolutional bounding box regression (Zhang et al., 2022). Retin a Net is an improvement over SSD that employs a feature pyramid network (FPN) to enhance feature extraction across different scales, improving object detection accuracy. Table 3 summarizes the RetinaNet model

Table 3 – Retina Net model summarization.

Type	advantage	Disadvantages	Domain
RetinaNet	• High accuracy. • Fast detection speed • Scalability • Simplicity	• Lower accuracy than some two-stage detectors • It can be less accurate for small objects • Requires more training data than some other models	• Image and Video Surveillance • Autonomous Vehicles. • Augmented Reality. • Robotics and Automation. • Drone Surveillance and Inspection. • Medical Imaging Analysis. • Sports Analytics and Performance Tracking. • Traffic Monitoring and Management. • Retail and Product Analysis.

8.2 Two-Stage Object Detection Algorithm

8.2.1 R-CNN (Region-based Convolutional Neural Network):

Is an innovative two-stage object detection system that locates and recognizes things in pictures or movies. 2014 saw its initial release. Region proposals were combined with convolutional neural networks (CNNs) to transform the object detection field, and it has since become the fundamental framework for object detection. It was thought to be a step toward developing more sophisticated algorithms (Bai, 2020) . Table 4 summarizes the R-CNN algorithm.

Table 4 - R-CNN algorithm summarization.

Type	advantage	Disadvantages	Domain
R-CNN	• Image features can be automatically extracted by using appropriate CNN Networks. • Regression analysis improves boundary box predictions and minimizes positioning errors.	• inadequate control over memory. Memory can run out on the machine if enough features are extracted. Can run out of memory. • More Complicated than most other existing algorithms at the time.	1. Security and Surveillance. 2. Medical Imaging. 3. Autonomous Vehicles. 4. Robotics. 5. Agricultural Applications. 6. Smart City Applications. 7. Wildlife Monitoring. 8. Satellite Imagery Analysis.

	● less error rate than CNNs that are conventional (Chun et al., 2020).	● More intricate than most other algorithms in use at the time. ● Slow operating speed makes it unsuitable for real-time applications (Chun et al., 2020).	9. Industrial Automation. 10. Material Defect Detection.

8.2.2 Fast R-CNN

The Region-based Convolutional Neural Network (R-CNN) is a foundational approach for object detection. Fast R-CNN addresses its computational efficiency issues. 2015 saw the introduction of Fast R-CNN (Girshick, 2015), which keeps R-CNN's outstanding accuracy while greatly accelerating object detection. Fast RCNN allows for improved speed and accuracy while addressing the shortcomings of CNN (Sukkar et al., 2021). Fast RCNN shares all network layer parameters, may operate momentarily as a feature cache without requiring disk space, employs multitask loss, has a one-stage training approach, and a higher MAP (mean average precision) than RCNN (Zhu et al., 2006), Table 5 summarize Fast R-CNN algorithm.

Table 5 - Fast R-CNN Algorithm Summarization.

Type	advantage	Disadvantages	Domain
Fast R-CNN	● multitasking training takes the role of the laborious stage-by-stage execution training approach. ● utilizes ROI pooling to satisfy requests for many scales (Mittal et al., 2020)	● As the primary selective search method, the time-consuming algorithm is inherently slow (Mittal et al., 2020; A. K. Shetty et al., 2021)	1. Security and Surveillance. 2. Traffic Monitoring and Analysis. 3. Medical Imaging Analysis. 4. Autonomous Vehicles. 5. Robotics and Automation. 6. Industrial Inspection and Quality Control. 7. Agricultural Applications. 8. Environmental Monitoring and Conservation. 9. Retail and Product Analysis. 10. Sports Analytics and Performance Evaluation.

8.2.3 Faster R-CNN

Faster R-CNN is a two-stage object identification algorithm that has revolutionized computer vision by enabling real-time object detection with high accuracy. Because of its durability and adaptability, it is a vital instrument for many different jobs and applications in several fields. Faster R-CNN (J. Wang et al., 2021), which was proposed three months after Fast R-CNN, further enhances the region-based CNN baseline. Slow and requiring the same amount of running time as the detection network, Fast R-CNN proposes ROI through selective search. A new RPN (region proposal network), a fully convolutional network that can effectively predict region proposals with a wide range of sizes and aspect ratios, takes its position in the Faster R-CNN; Table 6 summarizes the Faster R-CNN algorithm.

Table 6 - Faster R-CNN algorithm summarization.

Type	advantage	Disadvantages	Domain
Faster R-CNN	● The first CNN Family algorithm that yields good real-time performance. ●• Increased output accuracy and mAP.	● it takes a lot of time and expense to reshape the anticipated region proposals before estimating the bounding box's actual offsets.	1. Security and Surveillance. 2. Traffic Monitoring and Analysis. 3. Medical Imaging Analysis. 4. Autonomous Vehicles. 5. Robotics and Automation. 6. Industrial Inspection and Quality Control. 7. Agricultural Applications. 8. Environmental Monitoring and Conservation. 9. Retail and Product Analysis. 10. Sports Analytics and Performance Evaluation.

8.2.4 Mask R-CNN

Building on the work of Faster R-CNN, Mask R-CNN (Mask Region-based Convolutional Neural Network) is an object identification and instance segmentation system that can detect items and accurately define their borders down to the pixel level. This is accomplished by concurrently adding a branch for object mask (Region of Interest) prediction to the branch for bounding box recognition (He et al., 2020) . Table 7 summarizes this algorithm.

Table 7 - Mask R-CNN algorithm summarization.

Type	advantage	Disadvantages	Domain
Mask R-CNN	● The tendency is for all the operations classification, box regressions, and mask generation to operate in tandem to produce more effective outcomes.	● requires a lot of processing power because layers operate concurrently.	1. Autonomous Vehicles. 2. Medical Image Analysis. 3. Satellite Imagery Analysis. 4. Retail and Product Analysis. 5. Sports Analytics and Performance Evaluation. 6. Agricultural Applications. 7. Robotics and Automation. 8. Security and Surveillance. 9. Environmental Monitoring and Conservation. 10. Biomedical Research and Drug Discovery.

There are many algorithms for object detection, but the most famous is YOLO (You Only Look Once). Due to this algorithm's unique characteristics, have the advantages and disadvantages of YOLO.

9. YOLO algorithm

YOLO, an acronym for "You Only Look Once," is a group of real-time object identification algorithms renowned for their accuracy and speed. YOLO locates and identifies things in a scene with a single Stage, unlike conventional object identification algorithms that need two Stages across an image. Because of this, YOLO is perfect for real-time applications like video surveillance and self-driving automobiles. Yolo operates by first splitting the input image into a grid of cells. Each grid cell is subjected to a deep convolutional neural network (CNN) to extract the features. Predicts the bounding box coordinates for each item and the probability that each sort of object will exist in the cell. Yolo is a standard option for problems involving object detection. Yolo comes in various forms, each with unique advantages and disadvantages.

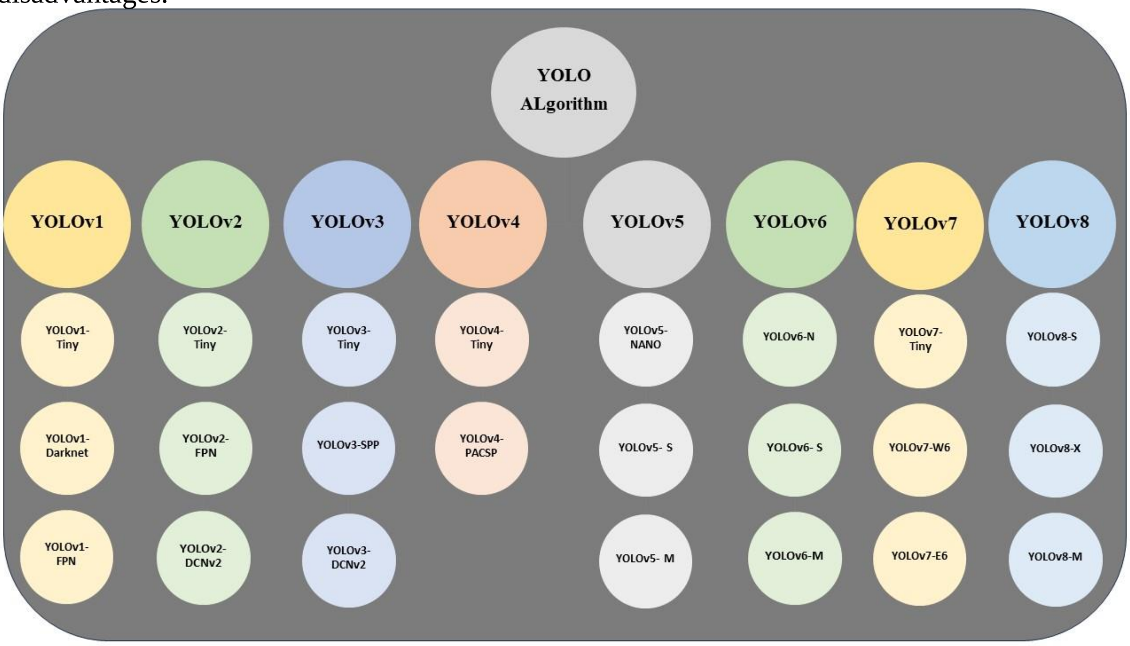

Fig. 3. A version of YOLO Algorithm.

.

Based on the above Figure, Table 8 Summarizes the Categories of the YOLO algorithm.

Table 8 - YOLO Type

Ref.	Vergne	Pros	Cons
(Cheng, 2020)	YOLOv1	YOLOv1 is used for real-time applications because of its speed. The general structure of it is simple. Reduced background error.	The bounding boxes that are projected for an item may not be accurate. Limited range of item sizes Compared to several other object identification methods, YOLOv1's accuracy is not as good.
(R. Li & Yang, 2018)	YOLOv2	Enhanced accuracy and Faster processing and Improved class separation from YOLOv1	complex architecture, Increased memory consumption, and Limited improvement in localization from YOLOv1
(Chun et al., 2020)	YOLOv3	YOLOv3 is a potent tool with processing rates that surpass 100 frames per second and improved accuracy due to architectural improvements.	Long Training Time, and The accuracy difference is significant compared to other models. In addition, his lack of accuracy in detecting small details
(Bochkovskiy et al., 2020)	YOLOv4	YOLOv4 boasts even faster processing than YOLOv3, reaching up to 150 frames per second; compared to the Yolov3, the Yolov4 accuracy is higher.	The yolov4 is more complicated than the rest of the previous versions, and this causes a longer training time. Also, its application to low-power devices is limited.
(Sukkar et al., 2021)	YOLOv5	Yolov5 has fantastic accuracy, often matching or exceeding earlier Yolo versions, fast speed (up to 140 frames per second), flexibility and simplicity of use.	Because of Yolo's complexity, there is little transparency, a substantial training period, and excessive training data processing.
(C. Li et al., 2022)	YOLOv6	YOLOv6 has more straightforward training procedures, improved noise and other obstacle tolerance, and faster processing for the Yolov6 up to 200 frames per second.	Yolov6 has some drawbacks, including more complexity than previous versions and a need for more resources for performance enhancements, which raises resource consumption.
(Optoelectronics, 2023)	YOLOv7	Yolov7 is designed to attain improved strength, precision, and speed. Its primary objective is to sustain the real-time processing rates of previous versions and maybe surpass two es per se, increasing efficiency, scalability, and simplicity of use.	Complicated design that often balances high performance against higher resource needs, as well as the need for more powerful devices at work.
(Huangfu & Li, 2023)	YOLOv8	With the simplicity of use and excellent efficiency, this version maintains or surpasses the real-time processing rates of previous versions at a rate of 300 frames per second by focusing more intently on the pertinent regions inside the images for accurate	complex architecture that often strikes a compromise between increased resource requirements and excellent performance.

10. Object Detection and Privacy Measures

One may use the concept of integrating technology to create an effective method for safeguarding individuals' information in digital images and videos. An idea proposed is to merge the detection algorithm with the protection technology for people and face protection. The Yolo algorithm, known for its effective results with Blur technology, could be utilized to implement this merger. By applying Blur technology to the face and body, individuals can be protected without compromising the image or video quality, or drawing attention to hidden information in the media. Combining person identification with privacy protections is a very effective approach for safeguarding the anonymity of persons shown in images or films. As technology continues to progress, we may anticipate the development of further tools and services that allow people to exchange material while safeguarding their privacy.

11. Famous Dataset

COCO, Pascal VOC, ImageNet, KITTY, and Visual Genome are key datasets for object identification. COCO offers 330,000 images and 80 item categories, while Pascal VOC is more minor but crucial. ImageNet provides categorization data, KITTY uses sensor data for reliable detection, and Visual Genome annotates activities like captioning and image comprehension. Human identification Included in object detection, person detection is essential to many applications, such as autonomous vehicles, video surveillance, and human-computer interaction. With different strengths and shortcomings, several datasets have been created to test and train human detection algorithms. Table 9 summarizes the famous dataset in human actions.

Table 9 - Famous Dataset In Human Actions.

Dataset	**Year**	**Description**	**Strengths**	**Weaknesses**	**images**
INRIA Person (Redmon et al., 2016)	2005	A popular dataset for computer vision pedestrian identification is the INRIA Person dataset. It has bounding boxes around pedestrians images in different settings and orientations.	Easily accessible, straightforward and practical, it is often used for benchmarking algorithms.	Small size and variety; primarily targets upright pedestrians in confined spaces.	2,400
MS COCO (Lin et al., 2014)	2018	The MS COCO (Microsoft Common Objects in Context) dataset is Utilized extensively for training and assessment purposes, object identification, and segmentation.	Rich annotations for many human traits and behaviours, large size and variety.	The main emphasis is not on human annotations; training models on this extensive dataset requires a significant computing investment.	330,000
Caltech Pedestrian (Brunetti et al., 2018; Yang et al., 2016)	2016	A larger and more challenging dataset contains varying lighting, weather conditions, and partial occlusions. This dataset offers scholars an extensive collection of images depicting the many and often tricky situations in real-world settings.	Greater variety and size, with difficult situations for reliable detection.	Still primarily focuses on pedestrians; it lacks annotations for clothes, gender, and age.	Over 300,000
Wider Face (Yang et al., 2016)	2021	A large-scale dataset created especially for assessing face identification systems is called the Wider Face dataset. It provides a wide range of images and annotations, which makes it a demanding and thorough standard for academics working in the subject.	The accepted benchmark for face identification methods is a widely used and popular dataset. It is openly accessible, allowing users to download and study without charge. Face landmark annotations are present in a subset of images, facilitating efficient training and evaluation of algorithms.	Limited Annotations and training algorithms on the large dataset require significant computational resources, Focus on Face Detection ,	393,073
ETH Zurich Human Pose (Yuan et al., 2019; Zurich et al., 2021)	2019	It is an invaluable resource for scientists and programmers working on human posture estimate, a crucial computer vision issue with applications across many domains.	Accurate body component tracking and identification are made possible by detailed position annotations.	Primarily centred on interior settings, with limited size and scene variation.	4,000

Table 10: summarises the previous steady in the INRIA Dataset.

Table 10 - Previous Steady INRIA Dataset.

Ref.	Year	author	aim	description	accuracy
(Yuan et al., 2019)(Wiratama & Sim, n.d.)	2019	Danni Yuan, Xiaoyan Zhu, Yaoru Mao, Binwen Zheng, Tao Wu	Pedestrian Detection	The INRIA Person Dataset was gathered to research the identification of upright individuals in images and videos. Currently one of the most widely used static pedestrian identification datasets, INRIA includes images from various sources, including personal digital image collections and a few Google images. The training and test datasets are the two components of the dataset. The test dataset has 1132 images with pedestrians and 453 images without, while the training dataset contains 2416 images with walkers and 1218 without. Each image in both datasets has been resized to 64 x 128 pixels.	97.3%.
(Matsumura & Hanazawa, 2019; Wiratama & Sim, n.d.)	2019	Wahyu Wiratama, Donggyu Sim	person detection in HEVC bitstream	This study uses the INRIA dataset, which contains positive and negative images and its annotations for training and testing.	MAP 0.68
(Matsumura & Hanazawa, 2019; Nakachi & Kiya, 2020)	2020	Matsumura1 and Akitoshi Hanazawa2	HUMAN DETECTION	Utilize the NICTA Pedestrian dataset with the INRIA Person dataset. The INRIA dataset includes 2,416 and 4,832 images of positive and negative samples, respectively, as training data. The NICTA dataset contains 6,500 images of negative samples and 3,000 images of positive samples, respectively, that are used as training data. The INRIA dataset contains 1,132 images representing both positive and negative samples, forming the test data.	95%
(Nakachi & Kiya, 2020)	2020	Takayuki Nakachi and Hitoshi Kiya	PRIVACY-PRESERVING PATTERN RECOGNITION WITH IMAGE COMPRESSION	in The INRIA Pedestrian dataset. The training set of the dataset comprises 614 positive images, which include 1237 pedestrians, whereas the testing set has 228 positive images, which contain 589 pedestrians. Each image in this collection has a resolution of 64 by 128 pixels.	80%
(Murthy & Hashmi, 2020)	2020	Chintakindi Balaram Murthy and Mohammad Farukh Hashmi	Real-Time Pedestrian Detection	The INRIA Pedestrian dataset has two positive images: the testing set consists of 228 positive images (which involves 589 persons), and the training set has 614 positive images (containing 1237 persons). This collection contains only images with resolutions of 64 × 128 pixels. The pedestrian images in this set include a variety of backgrounds, lighting conditions, levels of occlusion, human posture variations, and costumes.	79.86%
(Shao et al., 2021)	2021	Xiaoqiang Shao , Jinyang Wei1,Defeng	pedestrian detection	The INRIA pedestrian data collection, separated into training and test sets, is the most widely used static	improved by 6.3% and

		Guo,Runyang Zheng,Xinchao Nie ,Guowei Wang ,Yu Zhao		surveillance database. One thousand two hundred eighteen negative images and 614 positive ones made up the training set. Two hundred eighty-eight images make up the test set.	13.93% respectively
(Tahir et al., 2023)	2023	Azrina Tahir , Shamsul Kamal Ahmad Khalid, Lokman Mohd Fadzil	Identify children	collection of images and bounding boxes of children in various poses, environments, and lighting conditions. The dataset trained a YOLOv5 model to detect children in images.	Precision, recall and mAP_0.5 is 0.998, 0.995 and 0.995
(Weng, 2023)	2023	Guifan Weng	Real-time pedestrian recognition on low computational resources	The pedestrian dataset from INRIA is a standard one. In this image, the pedestrians are positioned in a very standard way. The majority include mostly people standing or strolling, taken from various perspectives with little deviance from a horizontal viewpoint.	95%

12. Conclusion

Privacy is essential to research because of the critical data of personal images [face and body]. Consequently, it is a complex idea that encompasses a range of factors: tracking and surveillance, identity theft and exploitation, and ethical and legal issues. So, the business landscape around face and body privacy is dynamic and diverse, offering various solutions to address the growing data collection and use concerns. This paper discusses the techniques used to hide people's information in the images. Therefore, it was concluded that Blur is the best possible way to use it because of its advantages, which represented the absence of image distortion and the inability to return the image to its previous position during a For the method. It is difficult to definitively say whether YOLO is the best method for human detection, as there are other object detection algorithms with their own strengths and weaknesses. However, after comparison, it became clear that YOLO has several advantages that make it a strong contender for the preference, especially in real-time applications. YOLO is a fast, accurate, simple object detection algorithm well-suited for real-time applications. Regarding the dataset and Based on the information discussed above, INRIA is a suitable choice for human detection tasks focusing on pedestrian identification in straightforward settings, particularly for initial model training or benchmarking. Focus characterizes it on pedestrians: Unlike COCO, which includes a more comprehensive range of object classes, INRIA specifically targets pedestrians, making it directly relevant for human detection tasks. In addition to Straightforward settings, the images primarily depict pedestrians in explicit poses and lighting conditions, offering a good starting point for training models before tackling more complex scenarios. Accessibility and ease of use: Compared to Wider Face, INRIA's smaller size and more straightforward annotations make it readily accessible and computationally efficient for initial training or benchmarking purposes.

References

Abed, R. M., Abdulmalek, H. W., Yaaqoob, L. A., Altaee, M. F., & Kamona, Z. K. (2023). Genetic Polymorphism of TLR5 and TLR6 in Iraqi Patients with Heart Failure Disease. *Iraqi Journal of Science*, *64*(4), 1662–1674. https://doi.org/10.24996/ijs.2023.64.4.9

Abood, Q. K. (2023). Predicting Age and Gender Using AlexNet. *TEM Journal*, *12*(1).

Adarsh, P., Rathi, P., & Kumar, M. (2020). YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. *2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020*, *March 2020*, 687–694. https://doi.org/10.1109/ICACCS48705.2020.9074315

Amer, F., Ali, M., & Al-Tamimi, M. S. H. (2022). Face mask detection methods and techniques: A review. *Int. J. Nonlinear Anal. Appl*, *13*(November 2021), 2008–6822. http://dx.doi.org/10.22075/ijnaa.2022.6166

Anushka, Arya, C., Tripathi, A., Singh, P., Diwakar, M., Sharma, K., & Pandey, H. (2021). Object Detection using Deep Learning: A Review. *Journal of Physics: Conference Series*, *1854*(1). https://doi.org/10.1088/1742-6596/1854/1/012012

Aribilola, I., Naveed, M., & Lee, B. (2022). *DEMIS : A Threat Model for Selectively Encrypted Visual Surveillance Data DEMIS : A Threat Model for Selectively Encrypted Visual Surveillance Data*. *October*, 1–17.

Bai, T. (2020). Analysis on Two-stage Object Detection based on Convolutional Neural Networkorks. *Proceedings - 2020 International Conference on Big Data and Artificial Intelligence and Software Engineering, ICBASE 2020*, 321–325. https://doi.org/10.1109/ICBASE51474.2020.00074

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. http://arxiv.org/abs/2004.10934

Brunetti, A., Buongiorno, D., Trotta, G. F., & Bevilacqua, V. (2018). Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing*, *300*, 17–33. https://doi.org/10.1016/j.neucom.2018.01.092

Patel, S. C., & Salot, P. (2021). Survey on Different Object Detection and Segmentation Methods. *Int. J. Innov. Sci. Res. Technol*, *6*(1), 608-611.

Cheng, R. (2020). A survey: Comparison between Convolutional Neural Network and YOLO in image identification. *Journal of Physics: Conference Series*, *1453*(1). https://doi.org/10.1088/1742-6596/1453/1/012139

Chettri, S. K., & Borah, B. (2015). Anonymizing classification data for preserving privacy. *Communications in Computer and Information Science*, *536*, 99–109. https://doi.org/10.1007/978-3-319-22915-7_10

Chun, L. Z., Dian, L., Zhi, J. Y., Jing, W., & Zhang, C. (2020). YOLOv3: Face detection in complex environments. *International Journal of Computational Intelligence Systems*, *13*(1), 1153–1160. https://doi.org/10.2991/ijcis.d.200805.002

Dhillon, A., & Verma, G. K. (2020). Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, *9*(2), 85–112. https://doi.org/10.1007/s13748-019-00203-0

Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, *2015 Inter*, 1440–1448. https://doi.org/10.1109/ICCV.2015.169

Gonzalez Dondo, D., Redolfi, J. A., Araguas, R. G., & Garcia, D. (2021). Application of Deep-Learning Methods to Real Time Face Mask Detection. *IEEE Latin America Transactions*, *19*(6), 994–1001. https://doi.org/10.1109/tla.2021.9451245

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2020). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(2), 386–397. https://doi.org/10.1109/TPAMI.2018.2844175

Hosny, K. M., Zaki, M. A., Hamza, H. M., Fouda, M. M., & Lashin, N. A. (2022). Privacy Protection in Surveillance Videos Using Block Scrambling-Based Encryption and DCNN-Based Face Detection. *IEEE Access*, *10*(September), 106750–106769. https://doi.org/10.1109/ACCESS.2022.3211657

Huangfu, Z., & Li, S. (2023). Lightweight You Only Look Once v8: An Upgraded You Only Look Once v8 Algorithm for Small Object Identification in Unmanned Aerial Vehicle Images. *Applied Sciences*, *13*(22), 12369. https://doi.org/10.3390/app132212369

Intelligence and Neuroscience, C. (2023). Retracted: Local Privacy Protection for Sensitive Areas in Multiface Images. *Computational Intelligence and Neuroscience*, *2023*, 1–1. https://doi.org/10.1155/2023/9898736

Khaled, F., & Al-Tamimi, M. S. H. (2021). Plagiarism Detection Methods and Tools: An Overview. *Iraqi Journal of Science*, *62*(8), 2771–2783. https://doi.org/10.24996/ijs.2021.62.8.30

Kouadra, I., Bouchra, M., Bekkouche, T., & Ziet, L. (2023). Encryption face area in color images using Chaotic Maps. *International Conference on Pioneer and Innovative Studies*, *1*, 501–506. https://doi.org/10.59287/icpis.880

Lander, K., Bruce, V., & Hill, H. (2001). Evaluating the Effectiveness of Pixelation and Blurring on Masking the Identity of Familiar Faces. *Applied Cognitive Psychology*, *15*(1), 101–116. https://doi.org/10.1002/1099-0720(200101/02)15:1<101::AID-ACP697>3.0.CO;2-7

Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Ke, Z., Li, Q., Cheng, M., Nie, W., Li, Y., Zhang, B., Liang, Y., Zhou, L., Xu, X., Chu, X., Wei, X., & Wei, X. (2022). *YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications*. http://arxiv.org/abs/2209.02976

Li, R., & Yang, J. (2018). Improved YOLOv2 Object Detection Model. *International Conference on Multimedia Computing and Systems -Proceedings, 2018-May*, 1–6. https://doi.org/10.1109/ICMCS.2018.8525895

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8693 LNCS*(PART 5), 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Matsumura, R. Y. O., & Hanazawa, A. (2019). Human detection using color contrast-based histograms of oriented gradients. *International Journal of Innovative Computing, Information and Control*, *15*(4), 1211–1222. https://doi.org/10.24507/ijicic.15.04.1211

Mittal, P., Singh, R., & Sharma, A. (2020). Deep learning-based object detection in low-altitude UAV datasets: A survey. *Image and Vision Computing*, *104*. https://doi.org/10.1016/j.imavis.2020.104046

Murthy, C. B., & Hashmi, M. F. (2020). Real time pedestrian detection using robust enhanced YOLOv3+. *Proceedings - 2020 21st International Arab Conference on Information Technology, ACIT 2020*. https://doi.org/10.1109/ACIT50332.2020.9300053

Nakachi, T., & Kiya, H. (2020). *Privacy-preserving Pattern Recognition with Image Compression*. 01–12. https://doi.org/10.5121/csit.2020.100201

Ning, C., Zhou, H., Song, Y., & Tang, J. (2017). Inception Single Shot MultiBox Detector for object detection. *2017 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2017, July*, 549–554. https://doi.org/10.1109/ICMEW.2017.8026312

Optoelectronics, S. (2023). *THE REAL TIME DETECTION OF TRAFFIC PARTICIPANTS USING YOLOV7 ALGORITHM Saldambe Rema N Marak, Vinodkumar Jadhav School of ECE, Dr. Vishwanath Karad MIT-WPU, Pune, India*. *42*(1), 1477–1485.

Padilla-López, J. R., Chaaraoui, A. A., & Flórez-Revuelta, F. (2015). Visual privacy protection methods: A survey. *Expert Systems with Applications*, *42*(9), 4177–4195. https://doi.org/10.1016/j.eswa.2015.01.041

Pulfer, E. (2019). Different Approaches to Blurring Digital Images and Their Effect on Facial Detection. *Computer Science and Computer Engineering*.

Rahma, M. M., & Salman, A. D. (2022). Heart Disease Classification-Based on the Best Machine Learning Model. *Iraqi Journal of Science*, *63*(9), 3966–3976. https://doi.org/10.24996/ijs.2022.63.9.28

Rakhmawati, L., Wirawan, & Suwadi. (2018). Image Privacy Protection Techniques: A Survey. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, *2018-Octob*(October), 76–80. https://doi.org/10.1109/TENCON.2018.8650339

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Shao, X., Wei, J., Guo, D., Zheng, R., Nie, X., Wang, G., & Zhao, Y. (2021). Pedestrian Detection Algorithm based on Improved Faster RCNN. *IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, *2021*, 1368–1372. https://doi.org/10.1109/IAEAC50856.2021.9390882

Shetty, A. K., Saha, I., Sanghvi, R. M., Save, S. A., & Patel, Y. J. (2021). A Review: Object Detection Models. *2021 6th International Conference for Convergence in Technology, I2CT 2021*, 1–8. https://doi.org/10.1109/I2CT51068.2021.9417895

Shetty, J., & Jogi, P. S. (2019). Study on different region-based object detection models applied to live video stream and images using deep learning. In *Lecture Notes in Computational Vision and Biomechanics* (Vol. 30). Springer International Publishing. https://doi.org/10.1007/978-3-030-00665-5_6

Shifa, A., Babar, M., Mamoona, I., Asghar, N., & Fleury, M. (n.d.). *Lightweight Human Skin Encryption for Public Safety in Real-time Surveillance Applications*.

Sukkar, M., Kumar, D., & Sindha, J. (2021). Real-Time Pedestrians Detection by YOLOv5. *2021 12th International Conference on Computing Communication and Networking Technologies, ICCCNT 2021, November*. https://doi.org/10.1109/ICCCNT51525.2021.9579808

Tahir, A., Ahmad Khalid, S. K., & Mohd Fadzil, L. (2023). Child Detection Model Using YOLOv5. *Journal of Soft Computing and Data Mining*, *4*(1), 72–81. https://doi.org/10.30880/jscdm.2023.04.01.007

Wang, J., Zhang, T., Cheng, Y., & Al-Nabhan, N. (2021). Deep learning for object detection: A survey. *Computer Systems Science and Engineering*, *38*(2), 165–182. https://doi.org/10.32604/CSSE.2021.017016

Wang, W. (2020). Detection of panoramic vision pedestrian based on deep learning. *Image and Vision Computing*, *103*, 103986. https://doi.org/10.1016/j.imavis.2020.103986

Wen, Y., Liu, B., Ding, M., Xie, R., & Song, L. (2022). IdentityDP: Differential private identification protection for face images. *Neurocomputing*, *501*, 197–211. https://doi.org/10.1016/j.neucom.2022.06.039

Weng, G. (2023). *Real-time pedestrian recognition on low computational resources*. http://arxiv.org/abs/2309.01353

Wiratama, W., & Sim, D. (n.d.). *A person detection in HEVC bitstream domain based on bits density feature and YOLOv3 framework*.

Xiao, X., & Tao, Y. (2006). Anatomy: Simple and effective privacy preservation. *VLDB 2006 - Proceedings of the 32nd International Conference on Very Large Data Bases*, 139–150.

Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). WIDER FACE: A face detection benchmark. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 5525–5533. https://doi.org/10.1109/CVPR.2016.596

Yuan, D., Zhu, X., Mao, Y., Zheng, B., & Wu, T. (2019). Privacy-Preserving Pedestrian Detection for Smart City with Edge Computing. *2019 11th International Conference on Wireless Communications and Signal Processing, WCSP 2019*, 1–6. https://doi.org/10.1109/WCSP.2019.8927923

Zhang, Y. M., Hsieh, J. W., Lee, C. C., & Fan, K. C. (2022). Sfpn: Synthetic Fpn for Object Detection. *Proceedings - International Conference on Image Processing, ICIP*, 1316–1320. https://doi.org/10.1109/ICIP46576.2022.9897517

Zhong, Y., & Deng, W. (2023). OPOM: Customized Invisible Cloak Towards Face Privacy Protection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(3), 3590–3603. https://doi.org/10.1109/TPAMI.2022.3175602

Zhu, Q., Avidan, S., Yeh, M. C., & Cheng, K. T. (2006). Fast human detection using a cascade of histograms of oriented gradients. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2*, 1491–1498. https://doi.org/10.1109/CVPR.2006.119

Zurich, E. T. H., Perception, M., Semester, S., Bucher, M., & Weiss, V. (2021). *3D Human Pose Estimation from RGB Images*. 5–8.