

Retraction

Retracted: Research on Fingerprint Security Based on Improved Yolo Algorithm

Mobile Information Systems

Received 31 October 2023; Accepted 31 October 2023; Published 1 November 2023

Copyright © 2023 Mobile Information Systems. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] X. Lu, "Research on Fingerprint Security Based on Improved Yolo Algorithm," *Mobile Information Systems*, vol. 2022, Article ID 5133471, 11 pages, 2022.

Research Article

Research on Fingerprint Security Based on Improved Yolo Algorithm

Xuefeng Lu 

School of Economics and Information, Zhejiang Tongji Vocational College of Science and Technology, Zhejiang, Hangzhou 311231, China

Correspondence should be addressed to Xuefeng Lu; z20220010802@zjtongji.edu.cn

Received 12 May 2022; Revised 21 June 2022; Accepted 21 July 2022; Published 7 September 2022

Academic Editor: Praveen Kumar Reddy Maddikunta

Copyright © 2022 Xuefeng Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The quantitative identification technology based on the statistical law of fingerprint features has become a new research difficulty and focus, and the automatic detection and classification of fingerprint features are the basis for realizing automatic fingerprint feature statistics. In this paper, a YOLO-based fingerprint feature detection method was proposed. First, a fingerprint feature dataset was established, which contained a total of 4,000 annotated fingerprint images; then, according to the characteristics of small size and dense distribution of fingerprint feature points, the YOLO network structure was improved, the original large target feature detection layer by 32-fold downsampling was deleted, and a new small feature fusion layer was added; the FPN, PAN, and SPP structures were used to achieve local and global feature extraction through multiple-scale fusion methods; finally, the SE channel attention mechanism module was added to effectively enhance the model robustness and detection ability of dense small objects. The experimental results show that compared with the improved FP-YOLO model of the original model, when the detection speed is basically unchanged, the mAP0.5 value is increased from 93.0% to 97.4%, and the weight is reduced by 3/4.

1. Introduction

Fingerprint is an extremely unique biological feature of the human body, and it is an important basis for public security organs to crack down on criminal crimes [1]. Based on the fingerprints left by the perpetrator at the crime scene, the identity of the perpetrator can be determined through fingerprint query, comparison, inspection, and identification, providing evidence support for judicial trials [2]. In order to solve the problem of identification standards, it was first proposed that 12 feature matching points must be met for the identification of personal identity, but there was no scientific explanation for the number of matching points [3]. Reference [4] points out that the probability of two fingerprints showing eight identical features but not belonging to the same person is about one in 10 trillion but does not directly answer the standard question of the number of matching points for fingerprint identification. The most famous wrong fingerprint case is the Madrid tram bombing in 2004. In this case, the police extracted an incomplete

fingerprint at the scene. Based on this fingerprint, the American police mistakenly identified others as the perpetrator [5]. In 2014, the Miami Police Department of the United States conducted statistics on fingerprint error identification, and the results showed that the false true rate of fingerprint identification was 3.0%, and the false error rate was 7.5%, which shows that the qualitative fingerprint identification conclusion is not completely reliable [6].

The emergence of judicial misjudged cases has made judicial trials put forward higher requirements for the accuracy, reliability, and scientificity of court evidence, and the inspection and evaluation of fingerprint evidence have also begun to shift to a likelihood ratio framework model with quantitative evaluation as the core [7]. The characteristics on which fingerprint identification is based include ridge ending, bifurcation, spur, crossover, island, independent ridge, and lake [8], where the endpoints can be subdivided into starting and ending points, and the bifurcation points can be subdivided into bifurcation points and junction points. However, the distribution of these fingerprint

features is not balanced, and the endpoints and bifurcation points are the most common, and their identification value is much lower than that of other types [9]. To achieve the quantification of fingerprint identification conclusions, it is necessary to count the distribution rules of various fingerprint feature points, but there is no statistical result based on big data at present. The existing fingerprint identification technology can only simplify the fingerprint features into point-line features with directions and cannot accurately identify the above seven types of features.

From the perspective of fingerprint identification and quantification, this paper studied the automatic detection method of fingerprint features based on YOLO [10], which lays a technical foundation for automatic statistics of fingerprint feature distribution. First, a fingerprint feature dataset was established, and then according to the characteristics of small size, dense distribution and overlapping of fingerprint feature targets, on the basis of the original model YOLO, the optimal detection layer was selected for many experiments, and the shallow features and deep semantic features were fused. And an attention mechanism module was added to achieve accurate identification and precise positioning of fingerprint features. In the future, AI object detection will become more and more important.

2. Related Work

Object detection methods can be divided into two categories: two stage and one stage. In the two-stage object detection, the objects are first localized and then classified, and the representative algorithms are R-CNN [11], Fast R-CNN, and R-FCN [12]. One-stage object detection regards object detection as a regression problem and performs localization and classification at the same time. Representative algorithms include YOLO (You Only Look Once) [13], SSD [14], and RetinaNet [15].

According to the literature [16], YOLO is based on the end-to-end network structure and simultaneously completes the two tasks of object detection and classification. The disadvantage is that the detection accuracy is not high. Later, YOLO900 and YOLOv3 versions appeared, which improved detection accuracy while maintaining high detection speed. YOLOv4 makes it possible for object detection to be trained on low-performance GPUs [17]. In the same year, the literature [18] proposed YOLOv5. The detection accuracy of this model is higher than the previous two stage object detection model, and the detection speed is fast. It can be well applied to embedded devices and mobile terminals for detection. Therefore, YOLOv5 has become the current one of the best performing network models for object detection.

With the development of artificial intelligence, deep learning is also gradually applied in the field of fingerprint recognition. Literature [19] proposed a Cap-FingerNet, a fingerprint pattern classification network based on capsule network. In literature [20], a deep convolutional neural network was used to learn and represent the local ridge structure of fingerprints, and a new fingerprint aggregation method was proposed to improve the retrieval efficiency. However, there are few studies related to the extraction and

detection of fingerprint feature points. Literature [21] proposed a network for fingerprint feature extraction under complex background based on CNN model but did not distinguish the types of features, which has certain limitations. In order to realize the quantitative evaluation of fingerprint identification, this paper applied the YOLOv5 algorithm to the field of fingerprint identification to realize the detection and positioning of five types of fingerprint features, which lays the foundation for the establishment of a data-based probability evaluation method for identification in the future.

3. Method Proposed in This Paper

3.1. Introduction to YOLOv5. YOLOv5 includes a total of four models with different depths and widths, which are distinguished according to the parameters in the C3 module. The network is YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x from shallow to deep. The object detection performance also increases in turn, and the network application is more flexible and changeable to meet the different detection needs. The YOLOv5 network structure is shown in Figure 1, consisting of Input, Backbone, Neck, and Output.

Input section: batch normalizes the input image dimensions. Using Mosaic data augmentation [22], the model training speed is improved by randomly rotating, flipping, and scaling four images, and then stitching them into one image as training data. And using the adaptive anchors calculation method, in each time training, automatically according to the dataset used, the clustering algorithm was used to calculate the best set of anchor box values.

Backbone network part: this part consists of Focus, quartic ConV, three C3 modules, and an SPP structure to extract feature maps of different sizes from the input image. Focus uses the slicing operation to crop the input image and then stack it and performs a downsampling operation on the input image. The C3 module is improved on the basis of the YOLOv4 backbone network CSP module and consists of three parts: conv, batchnorm, and SiLU activation functions. The SPP (Spatial Pyramid Pooling) module [23] is used for feature fusion. The structure of the SPP module is shown in Figure 2. Through pooling of three scales, the feature map of any size is fixed as a feature vector of the same length and transmitted to the fully connected layer to realize the fusion of multiple receptive fields.

Neck web section: the FPN (Feature Pyramid Networks) and PAN (Pyramid Attention Network) structures [24] are used to fuse feature maps at different levels. FPN transfers deep semantic features to shallow layers from bottom to top, thereby enhancing semantic representation at multiple scales. On the contrary, PAN transmits the localization information of the shallow layer to the deep layer from top to bottom and enhances the localization ability on multiple scales. These two structures jointly enhance the feature fusion ability of the neck network, obtain more contextual information, and reduce information loss.

Output section: after 8x downsampling, 16x downsampling, and 32x downsampling, a total of three feature maps are generated at the network output. The smaller the

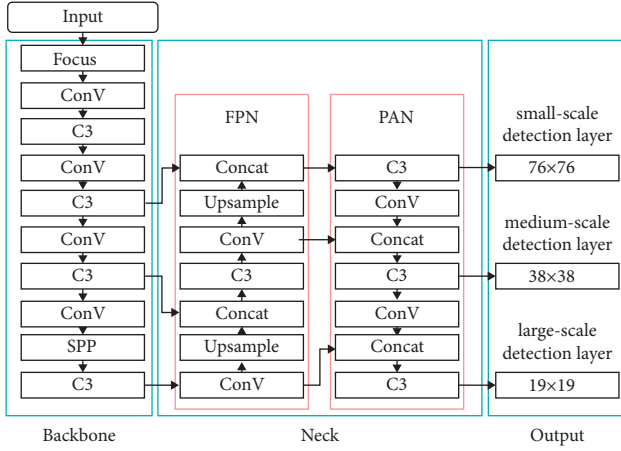


FIGURE 1: YOLOv5 network structure diagram.

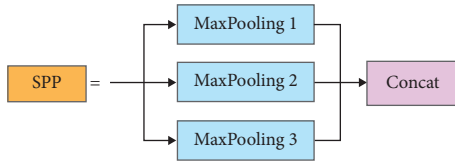


FIGURE 2: SPP network structure diagram.

feature map, the larger the image area corresponding to each grid cell. The output of 19×19 is suitable for detecting large-sized objects, while 76×76 is suitable for detecting small-sized objects. In YOLOv5, the CIOU_Loss is used as the loss function of the Bounding box [25]. Based on these new feature maps, the network output performs detection and classification.

3.2. Attention Mechanism. In recent years, attention mechanism has been widely used in various deep learning tasks such as computer vision and natural language processing. It has made many breakthroughs and has become a hot spot in neural network research. The most representatives are SE (Squeeze-and-Excitation) attention mechanism module and the proposed CBAM (Convolutional Block Attention Module) attention mechanism module.

The SE module is an attention mechanism for channels, including squeeze and excitation. The squeeze part is to perform global average pooling on the input. When the input size is $W \times H \times C$, the feature map is pooled and the output is a $1 \times 1 \times C$ vector. The excitation part is composed of two fully connected layers. In order to reduce the number of channels and the amount of parameters, the SERatio scaling parameter is added. The number of neurons in the first fully connected layer is $C \times \text{SERatio}$, and the output is $1 \times 1 \times C \times \text{SERatio}$. The number of neurons in the second fully connected layer is C , and the output is $1 \times 1 \times C$. The scale operation is to multiply the weights of each channel calculated by the SE module and the corresponding channels of the original input $W \times H \times C$, respectively, to rescale the

original features in the channel dimension. The SE module structure is shown in Figure 3(a).

The CBAM module extracts meaningful attention features from the Channel and Spatial dimensions successively. CBAM channel attention is roughly the same as SE module, the difference is that CBAM adopts max pooling and global average pooling in channel squeeze. The CBAM spatial attention structure is shown in Figure 3(b). The output of the channel attention module is used as the input of the spatial attention module $W \times H \times C$, again using max pooling and global average pooling to obtain two $W \times H \times 1$ feature maps. After 7×7 convolution kernel convolution and scale operation, the feature map adjusted by the dual attention mechanism is obtained.

The formulas of the channel attention mechanism M_c and the spatial attention mechanism M_s of the CBAM module are as follows:

$$M_c(F) = \sigma \cdot (MLP(\cdot \text{AvgPool}(F))) \cdot + \cdot MLP(\text{MaxPool}(F)) \cdot \cdot F',$$

$$M_s(F') = \sigma \cdot \left(f \cdot 7 * 7 \left(\left[\cdot \text{Avgpool}(F') ; \cdot \text{MaxPool}(F') \right] \right) \right),$$
(1)

where F represents the input feature map of the channel attention mechanism; σ represents the Sigmoid activation function; MLP represents the parameters of the multilayer perceptron; F' represents the output of the channel attention, which is also the input of the spatial attention; $f \cdot 7 \times 7$ represents that the convolutional layer uses a 7×7 convolution kernel.

3.3. YOLOv5 Model Improvements. The features in fingerprint images are small in size, large in number, densely distributed, and often overlapped, so it is not ideal to directly apply object detection methods such as YOLOv5 to detect features. In order to solve these problems, this paper makes three improvements to the original network structure of YOLOv5: (1) delete the 32-fold downsampling large-size feature fusion layer and add a 4-fold downsampling small feature fusion layer; (2) migrate the FPN and PAN structures to the pruned network and select the appropriate SPP pooling kernel parameters; (3) add the SE attention mechanism module, as shown in Figure 4.

First, to improve the performance of YOLOv5 in detecting small-sized objects, a new tiny feature fusion layer was added. The fusion layer was output by quadruple downsampling of the backbone network and then fused with the eightfold downsampling feature map to generate a feature map with a size of 152×152 . The segmented grid was denser, which is helpful for small object detection and identification. Because there were no large-size features in the fingerprint image, the original structure 32 times downsampling large-size feature fusion layer and its corresponding backbone and neck network structure (dotted line part) were deleted, which greatly reduced the algorithm complexity and the amount of parameters of the model.

Second, in order to improve the detection performance of the model for dense overlapping objects, the FPN and PAN structures were transferred to the pruned network. The

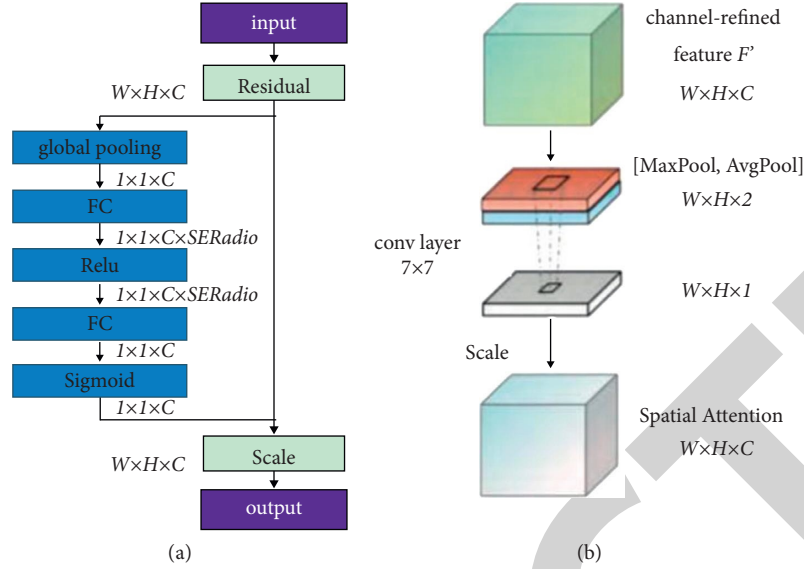


FIGURE 3: Attention mechanism module structure. (a) SE module structure; (b) spatial attention mechanism structure in CBAM module.

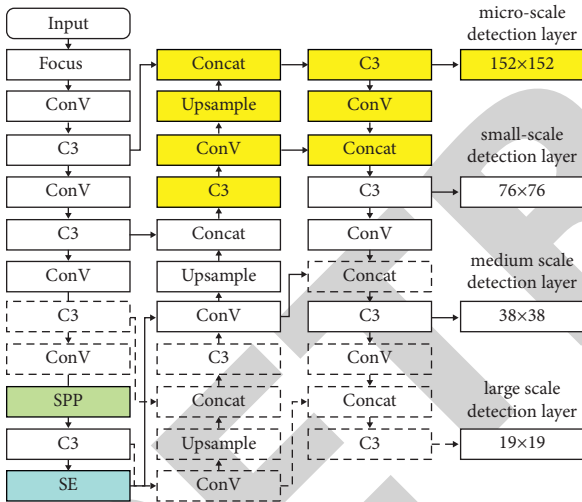


FIGURE 4: Improved network structure.

FPN structure transfers strong semantic features from top to bottom, and the PAN structure is a fusion from bottom to top, enhancing the localization ability at multiple scales. Continuing to use the SPP module, the feature maps were max-pooled from three different scales, which effectively increased the receiving range of the backbone features and realized the fusion of multiple receptive fields, which was beneficial for detecting large differences in object sizes and overlapping situations.

Finally, due to the uneven distribution of the number of fingerprint features, an SE attention mechanism module was added between the backbone and the neck network. Each feature map in channel attention mechanism represents a feature channel, which helps to filter out more meaningful features of the original image, focusing on feature channel weight assignment.

4. Experiment and Result Analysis

4.1. Fingerprint Dataset Construction

4.1.1. Collection of Datasets. The quality of the dataset has a huge impact on the design and training of object detection algorithms. Due to the low image quality of the current open-source fingerprint dataset, some fingerprints are incomplete. The experimental data come from the fingerprint images of the police in the actual cases, with a size of 680×680 and a resolution of 600 dpi, with a total of 500 images. The complete lines of the fingerprint are clear, which is conducive to subsequent preprocessing and reducing errors.

4.1.2. Dataset Preprocessing. Directly using deep learning for dimensionality reduction and feature extraction of fingerprint images will greatly affect the experimental accuracy, so fingerprint images need to be preprocessed. At present, the preprocessing method for fingerprints has been relatively perfect. The preprocessing steps in this paper are background separation, calculation of local ridge direction, ridge enhancement and binarization, as shown in Figure 5.

4.1.3. Dataset Labeling. Use Labelling software to make detection labels, manually label the features of 500 fingerprint images after preprocessing, and frame the full picture of feature points as accurately as possible to avoid framing irrelevant lines. The label format is set to the format of the PASCAL VOC dataset, and five types of features are labeled. The label names are bifurcation (label 0), spur (label 1), independent ridge (label 2), lake (label 3), and crossover (label 4). The image annotation is shown in Figure 6.

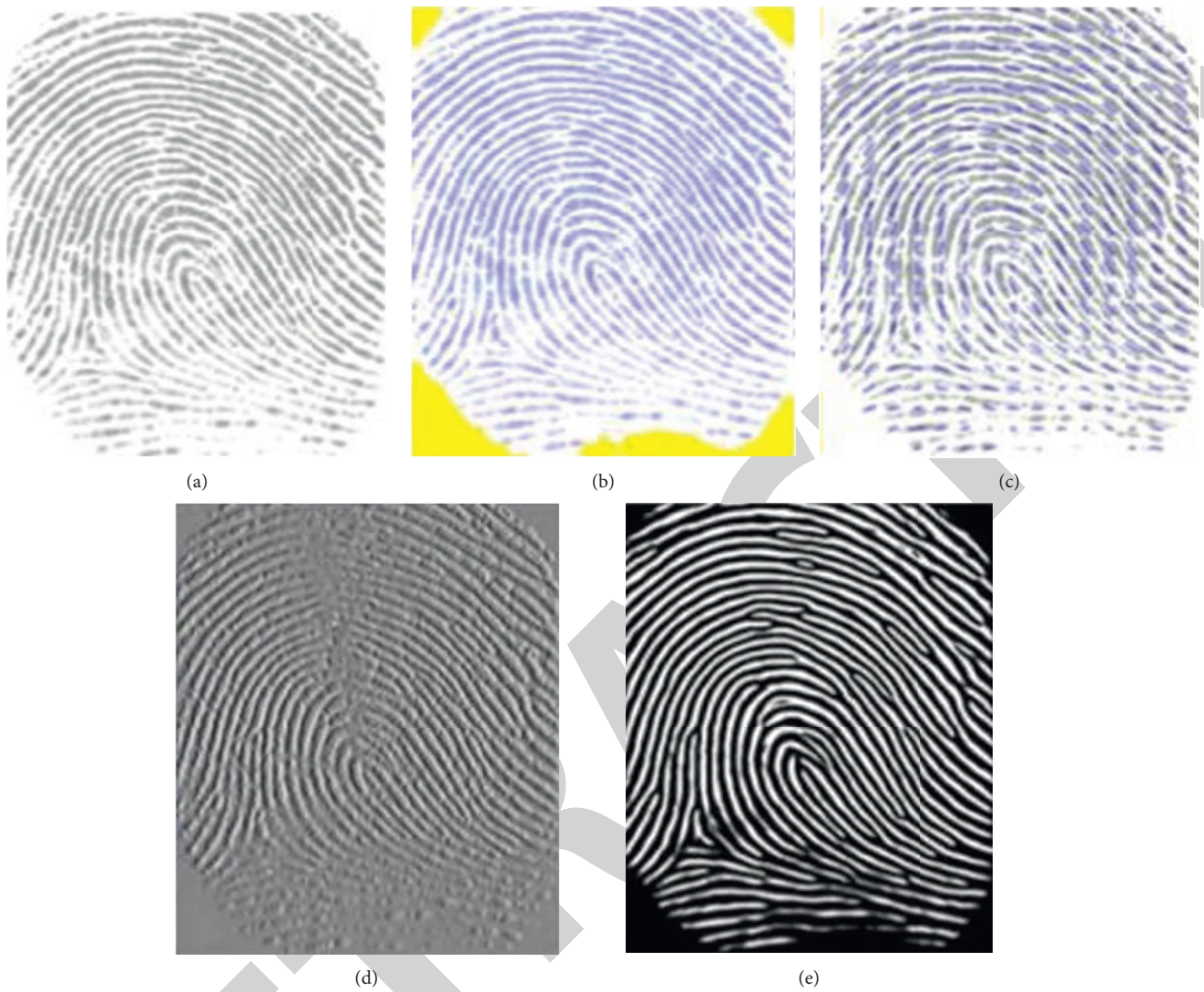


FIGURE 5: Effect of fingerprint image preprocessing. (a) Original image; (b) background segmentation; (c) local ridge orientation; (d) ridge enhancement; (e) binarization.

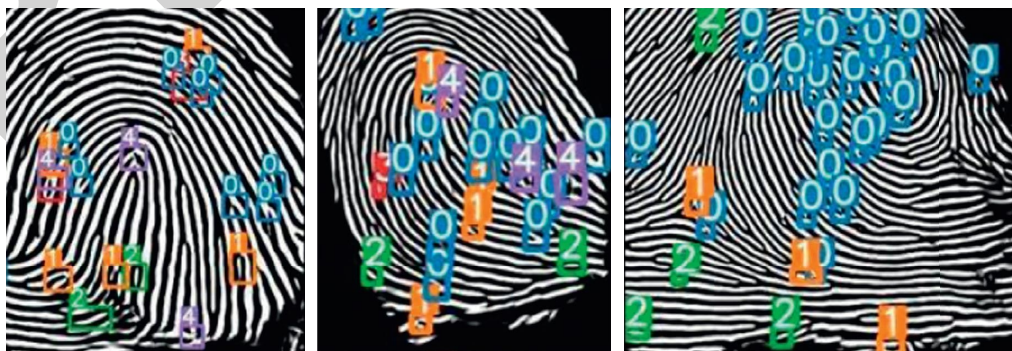


FIGURE 6: Image annotation.



FIGURE 7: Dataset enhancement effect.

TABLE 1: Distribution of various objects in training set and test set.

Object	Bifurcation	Crossover	Independent ridge	Lake	Spur
Training set	74985	2451	17811	3584	8948
Testing set	8319	285	1981	408	996

4.2. *Experimental Environment and Parameter Configuration.* The experimental platform operating system is Intel (R) Xeon (R) CPU E5-1650 v3 @ 3.50 GHz with 16 GB memory, GPU is NVIDIA GeForce GTX2080Ti with 11 GB video memory. The software configuration is Windows 10, CUDA1.2GPU parallel computing library, and the deep learning framework is Pytorch1.9.0 version.

In the training and testing of this experiment, the picture is set to 640×640 JPG format, the Batch_size is set to 36, the whole training process is 400 epochs, and the average precision mAP0.5; mAP0.5: 0.95%, weight, and actual detection speed FPS are used as model evaluation indexes for comparison.

Since the preset hyperparameters of the YOLOv5 model were optimized based on the COCO dataset, they were not universal. Therefore, hyperparameter evolution was used to obtain hyperparameter values that are more suitable for this dataset. The hyperparameter evolution algorithm used the genetic algorithm to adjust and optimize the hyperparameters according to the evaluation indicators and repeated the training process for 300 generations to obtain an initial learning rate (lr0) of 0.0128; cyclic learning rate (lrf) of 0.256; SGD learning rate momentum of 0.905.

Considering the insufficient samples of the dataset and the time cost of manual annotation, in order to increase the diversity of samples, data enhancement was performed on the dataset. It is known that the inversion and rotation of the fingerprint image do not affect the type and number of feature points. Flip and amplify each image in the vertical direction to double, and then rotate 90° , 180° and 270° clockwise, respectively. The amplification effect is shown in Figure 7.

After the above steps, the dataset was expanded eight-fold, with a total of 4,000 images and a total of 119,768 labels. The number of five types of features and their distribution in the training set and test set are shown in Table 1.

Mosaic data enhancement is a highlight of YOLOv5, by randomly rotating, flipping, and scaling four images, and then stitching them into one image as training data. For this dataset, some pictures have been rotated and flipped during data enhancement. Using Mosaic data enhancement will cause overfitting, and the method of scaling and splicing is

TABLE 2: YOLOv5 algorithm performance comparison table.

Model	mAP0.5 (%)	mAP0.5:0.95 (%)	Weight (M)
YOLOv5s	93.0	57.1	14.8
YOLOv5m	90.3	56.6	42.9
YOLOv5l	85.7	52.1	93.8
YOLOv5x	83.8	53.5	175.2
YOLOv5s_A	94.7	57.6	14.8

not conducive to the detection of small objects on fingerprint feature points. Therefore, Mosaic data enhancement is not used in this paper.

4.3. Experimental Results and Analysis

4.3.1. *Comparison of YOLOv5 Basic Model Results.* The YOLOv5 object detection network structure has four models: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, and the network depth and width increase successively. In this experiment, the model evolved from YOLOv5s for the hyperparameters of this dataset is named YOLOv5s_A. The YOLOv5 basic model algorithm is compared and tested using the fingerprint feature dataset constructed and annotated by ourselves. The indicators are shown in Table 2.

The structural complexity of YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x increases in turn. The more parameters, the greater the weight of the corresponding training-generated model, and the longer the training time. The test results on public datasets show that the more complex the YOLOv5 structure and the deeper the depth, the better the detection effect. However, different results have appeared in the object detection of fingerprint feature points. After analysis, too deep network and too many convolution operations are not suitable for detecting fine and small fingerprint feature points. It can be seen from Table 2 that compared with YOLOv5s, the mAP0.5 of the YOLOv5s_A model after hyperparameter evolution is increased by 1.7%, which effectively enhances the detection performance. Subsequent experiments used the values after hyperparameter evolution.

4.3.2. *Influence of Network Detection Layer on Detection Performance.* The original structure of YOLOv5s has three detection layers, which have undergone 8 times, 16 times, and 32-fold downsampling of the backbone network, respectively. The output feature map size corresponds to 76×76 , 38×38 , and 19×19 , respectively, realizing small, medium, and large-scale object detection. In order to explore the impact of different deep network detection layers on the

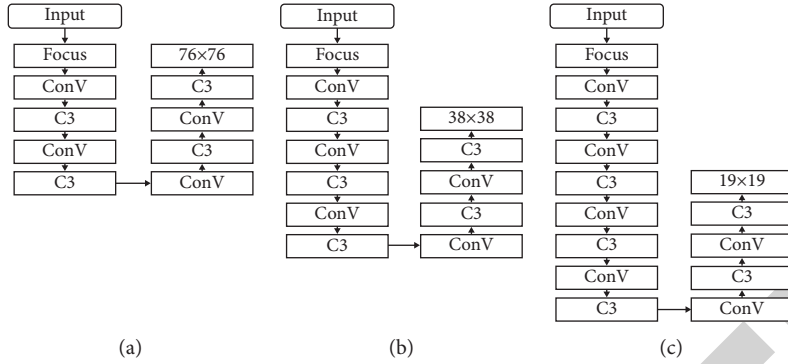


FIGURE 8: Model structure diagram. (a) YOLOv5s_8; (b) YOLOv5s_16; (c) YOLOv5s_32.

detection performance, this experiment selected 8-fold, 16-fold, and 32-fold downsampling layers as the detection layers from shallow to deep, and the distribution corresponds to three models of YOLOv5s_8, YOLOv5s_16, YOLOv5s_32, and their structures are shown in Figure 8.

Through experiments, it can be seen that the YOLOv5s_8 model with 8-fold downsampling as the detection layer has the best performance, with mAP0.5 being 67.8%, followed by 16-fold downsampling, and 32-fold downsampling accuracy is significantly reduced, as shown in Table 3. After analysis, when the number of downsampling layers in the backbone network is shallow, the lower spatial features are mainly extracted, which is helpful for the detection of small objects. The fingerprint feature points are small in size and densely distributed. If the backbone network is too deep, deeper semantic features will be obtained. Otherwise, detailed features will be lost, and the missed detection rate of small objects will be greatly improved, resulting in a significant drop in accuracy and an increase in the amount of experimental calculation. Therefore, the 32-fold downsampling layer is deleted in this experiment, and the 8-fold and 16-fold downsampling detection layers are retained.

4.3.3. Influence of Feature Fusion on Detection Performance.

After determining the depth of the Backbone network, on the basis of YOLOv5_16, Backbone adds the SPP pyramid pooling module, and Neck uses the FPN and PAN structure to fuse the features of the 8-fold and 16-fold downsampling layers, naming the model YOLOv5s_B. In order to explore the optimal pooling effect of SPP, four common SPP pooling kernels are tested in this paper: (3, 5, 7), (5, 7, 9), (7, 9, 13) and (9, 11, 13), named YOLOv5s_B_a, YOLOv5s_B_b, YOLOv5s_B_c, and YOLOv5s_B_d, respectively. Its structure is shown in Figure 9.

It can be seen from the experiments that YOLOv5s_B_c with SPP pooling kernel of (7, 9, 13) has the best detection performance, mAP0.5 is 93.7%, which is 24.3% higher than that of YOLOv5s_16, and the model weight is only increased by 0.7M, as shown in Table 4. As for why the model detection performance will be so significantly improved after feature fusion, after research, we find this is because the SPP module uses three different scales of maximum pooling for

TABLE 3: Performance comparison of different depth detection layers.

Model	mAP0.5 (%)	mAP0.5:0.95 (%)	Weight (M)
YOLOv5s_8	67.8	33.0	1.4
YOLOv5s_16	66.9	31.8	2.7
YOLOv5s_32	31.3	13.0	8.1

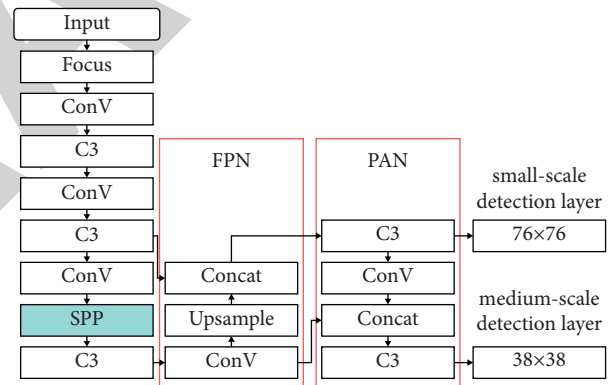


FIGURE 9: YOLOv5s_B network structure.

TABLE 4: Comparison of performance indicators of different SPP pooling kernels.

Model	mAP0.5 (%)	mAP0.5:0.95 (%)	Weight (M)
YOLOv5s_16	66.9	31.8	2.7
YOLOv5s_B_a	90.8	54.1	3.4
YOLOv5s_B_b	92.4	55.6	3.4
YOLOv5s_B_c	93.7	57.7	3.4
YOLOv5s_B_d	91.2	54.9	3.4

processing, which more effectively increases the receiving range of backbone features and realizes the fusion of multiple receptive fields. This effectively compensates for the loss of deep semantic information lost by 32-fold downsampling, and multilevel feature extraction also enhances the robustness of the network. The FPN structure transfers strong semantic features from the top feature map to the lower feature map for prediction at multiple scales. At the

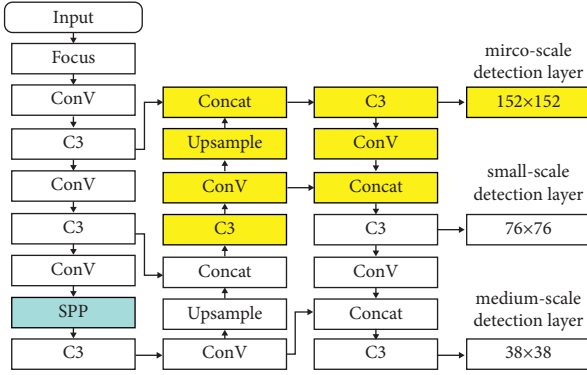


FIGURE 10: YOLOv5s_C network structure.

TABLE 5: Performance comparison after adding microscale detection layer.

Model	mAP0.5 (%)	mAP0.5:0.95 (%)	Weight (M)
YOLOv5s_B_c	93.7	57.7	3.4
YOLOv5s_C	95.2	58.7	4.0

same time, the PAN structure is a bottom-to-top fusion, transferring strong localization features from lower feature maps to higher feature maps, enhancing localization capabilities at multiple scales. The extraction of local and global features is achieved through the fusion of multiple scales, which enhances the expressive ability of the network and is conducive to detecting large differences in object sizes and overlapping features. Later experiments will be based on the YOLOv5s_B_c model to improve.

4.3.4. Influence of Adding Microscale Detection Layer on Detection Performance. Due to the small size and dense distribution of fingerprint feature points, and the small-scale 76×76 detection layer of YOLOv5 is not suitable for fingerprint feature points, this paper tries to add a new microscale detection layer. The detection layer is four-fold downsampled, and the model is named YOLOv5s_C. Its structure is shown in Figure 10.

It can be seen from the experimental results that after adding the microscale detection layer, mAP0.5 is 95.2%, and mAP0.5:0.95 is increased by two percentage points, which inevitably leads to a slight increase in the model weight, as shown in Table 5. The new four-fold downsampling detection layer makes the detection network structure more extensive and detailed and generates feature maps by extracting lower spatial features and fusing them with deep semantic features, which is suitable for detecting tiny, overlapping targets in fingerprint images.

4.3.5. Effect of Adding Attention Mechanism on Detection Performance. Attention mechanism is a method to force the learning process to focus on important channels and regions of the input object by adjusting different weights. In order to

explore whether adding an attention mechanism can optimize the detection performance, the attention mechanism modules CBAM and SE were added between the Backbone and Neck of YOLOv5s_C in turn, and the models were named YOLOv5s_CBAM and YOLOv5s_SE, respectively. The structure is shown in Figure 11(a).

The experimental results are shown in Table 6. Adding SE attention module has the best detection effect, mAP0.5 is 97.3% increased by 1.4%, mAP0.5:0.95 is increased by 3.2%, and the weight is only slightly increased. As for channel attention mechanism, each feature map represents a feature channel, which helps to filter out the meaningful features of the original image. In the spatial attention mechanism, one pixel in each feature map represents the feature of a certain area in the original image, which helps to train the network to pay attention to the feature of which area in the original image. SE only focuses on the channel weight assignment. CBAM considers both the importance of pixels in different channels and the importance of pixels in different positions of the same channel. Why is the accuracy not as good as SE attention mechanism? After analysis, there are two reasons: first, as shown in Figures 11(b) and 11(c), the fingerprint feature points are more distributed in the center and upper half of the fingerprint image. After adding the spatial attention mechanism, the network pays more attention to the center and upper part of the image, resulting in a decrease in attention to other spaces of the image and missing objects in the rest of the space. Second, it can be seen from Figure 11(a) that after the image is pooled by the SPP module with a pooling kernel of (5, 9, 13), the attention mechanism is passed to the spatial attention mechanism using a 7×7 convolution kernel for convolution. The channel compression methods of operation, max pooling and global average pooling, redundant repeated convolution of SPP, and CBAM modules cause the loss of useful information of the feature map, thus affecting the detection effect. Therefore, for fingerprint feature detection, the single channel attention mechanism SE module has better detection effect than the CBAM module.

It can be seen from the experimental data that YOLOv5s_SE performs the best, and the model is named YOLOv5s_Fingerprints Identification (hereinafter referred to as FP-YOLO).

4.3.6. Performance Comparison between the Improved Model and Other Detection Models. In this paper, FP-YOLO is compared with several algorithms with excellent performance at present, and the performance indicators are compared in Table 7.

It can be seen from Table 7 that for fingerprint feature point recognition, the accuracy rate, recall rate, and mAP0.5 of the FP-YOLO model proposed in this paper are better than the classical algorithms of SSD, YOLOv4, and YOLOv5s. The FPS of FP-YOLO is slightly lower than that of YOLOv5s. But the model training weight is only 4.1M, which

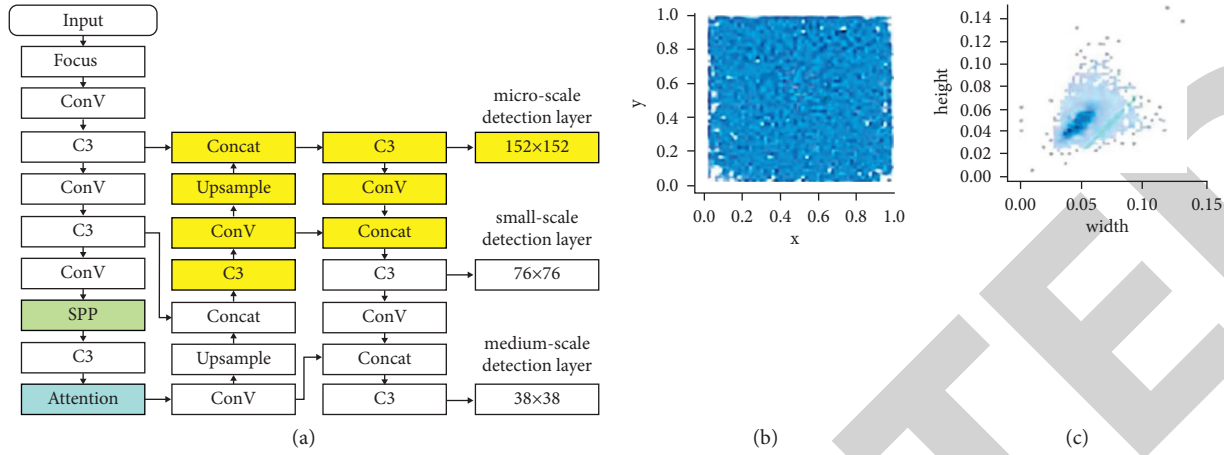


FIGURE 11: (a) YOLOv5s_attention network structure; (b) object location distribution of the dataset; (c) object size ratio of the dataset.

TABLE 6: Performance comparison of adding attention mechanism.

Model	mAP0.5 (%)	mAP0.5:0.95 (%)	Weight (M)
YOLOv5_C	95.2	58.7	4.0
YOLOv5_CBAM	94.7	58.5	4.1
YOLOv5_SE	97.4	61.9	4.1

TABLE 7: Performance comparison of different detection algorithms.

Model	Prediction (%)	Recall (%)	mAP0.5 (%)	Weight (M)	FPS/f/S-1
SSD	10.8	59.9	21.2	94.9	21.8
YOLOv4	80.4	46.6	60.6	250.3	32.2
YOLOv5s	93.6	87.0	93.0	14.8	27.3
FP-YOLO	96.5	94.0	97.	4.1	26.6

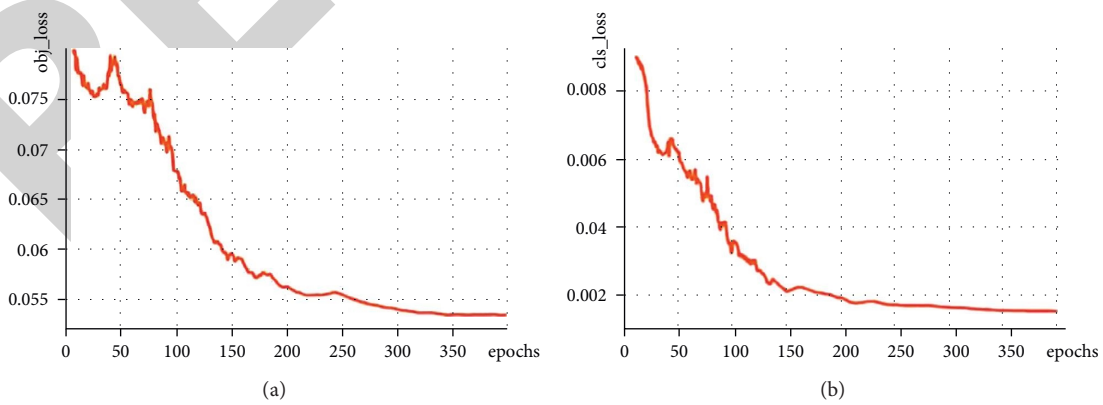


FIGURE 12: Training situation of FP-YOLO (a) Object loss change curve; (b) classification loss change curve.

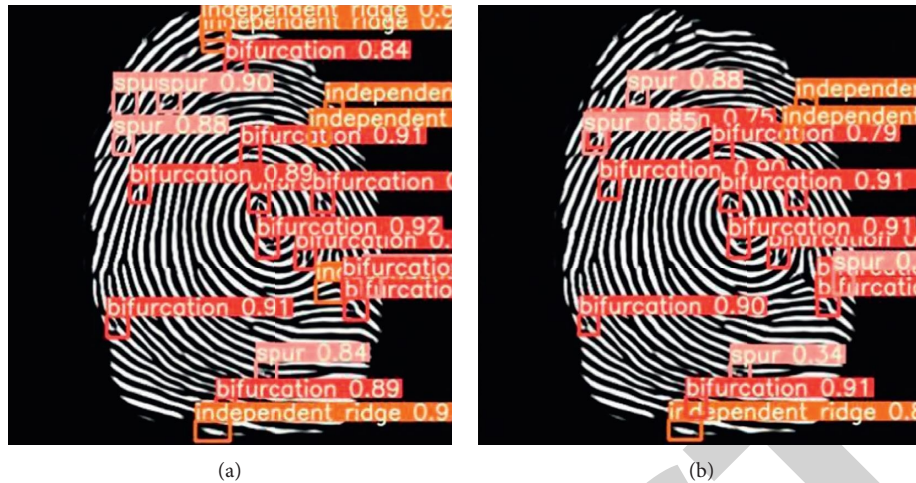


FIGURE 13: Comparison of test results. (a) FP-YOLO; (b) YOLOv5s.

is equivalent to a quarter of the YOLOv5s model. FP-YOLO achieves better performance in a comprehensive performance.

4.3.7. Improved Model Effect and Performance Evaluation. The object detection loss (obj_loss) curve and the classification loss (cls_loss) curve of the improved FP-YOLO base model at 400 rounds of training are shown in Figure 12.

The detection results of the FP-YOLO and YOLOv5s models on the validation set are shown in Figure 13. The improved model detection has more comprehensive objects, lower missed detection and false detection rates, and higher confidence in the detection results.

5. Conclusion

In view of the current difficulties in quantitative evaluation of fingerprint identification, in order to realize the mathematical statistics of five types of fingerprint feature points, this paper established five types of fingerprint feature datasets, based on this dataset, conducted training and comparison experiments, and improved the YOLOv5s algorithm, deleting 32-fold downsampling detection layer, adding 4-fold downsampling tiny feature fusion layer to effectively obtain more tiny feature information of fingerprint images. Using FPN, PAN, and SPP structures, local and global feature extraction is achieved through the fusion of multiple scales, which enhances the expressiveness of the network. In addition, the SE channel attention mechanism module is added, which reduces the interference of useless feature information to the model and enhances the channel weight of important features, thereby improving the detection effect. The experimental results show that the mAP0.5 accuracy of the FP-YOLO algorithm proposed in this paper reaches 97.4%, and the model weight is reduced by 72.3% under the condition that the detection speed is basically unchanged, which effectively increases the robustness of the model and the detection of dense small objects, realizing the accurate identification and positioning of five

types of fingerprint feature points. Ways to use artificial intelligence will become increasingly important in the future.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- [1] C. Van Der Walt, "Peace is not the absence of war but the presence of a relationship founded by God-(shalom) in Isaiah and Micah," *In die Skriflig/In Luce Verbi*, vol. 55, no. 2, pp. 1–8, 2021.
- [2] M. Rubab and D. Hassan, "Dehff – a hybrid approach based on distinctively encoded fingerprint features for live fingerprint detection," *Biomedical Signal Processing and Control*, vol. 75, 2022.
- [3] A. Husseis, J. Liu-Jimenez, and R. Sanchez-Reillo, "Fingerprint presentation attack detection utilizing spatio-temporal features," *Sensors*, vol. 21, no. 6, p. 2059, 2021.
- [4] S. Pankanti, S. Prabhakar, A. Jain, and A. K. Jain, "On the individuality of fingerprints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1010–1025, 2002.
- [5] Y. Zhang, C. Gao, S. Pan, Z. Li, Y. Xu, and H. Qiu, "A score-level fusion of fingerprint matching with fingerprint liveness detection," *IEEE Access*, vol. 8, pp. 183391–183400, 2020.
- [6] N. Kumar, K. Mahadevan, G. Nagaraju, and K. M. Mahadevan, "Development and detection of level II and III features of latent fingerprints using highly sensitive AIE based coumarin fluorescent derivative," *Journal of Science: Advanced Materials and Devices*, vol. 5, pp. 520–526, 2020.

- [7] G. Horsman, "The different types of reports produced in digital forensic investigations," *Science & Justice*, vol. 61, no. 5, pp. 627–634, 2021.
- [8] L. Ghiani, A. Hadid, G. L. Marcialis, F. Roli, and G. L. Marcialis, "Fingerprint liveness detection using local texture features," *IET Biometrics*, vol. 6, no. 3, pp. 224–231, 2017.
- [9] K. J. Bergen and G. C. Beroza, "Earthquake fingerprints: extracting waveform features for similarity-based earthquake detection," *Pure and Applied Geophysics*, vol. 176, no. 3, pp. 1037–1059, 2019.
- [10] R. Agrawal, A. S. Jalal, and K. V. Arya, "Fake fingerprint liveness detection based on micro and macro features," *IJBM*, vol. 11, no. 2, p. 177, 2019.
- [11] P. M. Blok, G. Kootstra, H. E. Elghor, B. Diallo, F. K. Van Evert, and E. J. Van Henten, "Active learning with MaskAL reduces annotation effort for training Mask R-CNN on a broccoli dataset with visually similar classes," *Computers and Electronics in Agriculture*, vol. 197, p. 197, 2022.
- [12] K. R. Akshatha, A. K. Karunakar, S. B. Shenoy, A. K. Pai, N. H. Nagaraj, and S. S. Rohatgi, "Human detection in aerial thermal images using faster R-CNN and SSD algorithms," *Electronics*, vol. 11, p. 11, 2022.
- [13] X. Xueqian, C. Xinqiang, W. Bing, Z. Wang, and J. Zhen, "Exploiting high-fidelity kinematic information from port surveillance videos via a YOLO-based framework," *Ocean & Coastal Management*, vol. 222, 2022.
- [14] Y. Miao, S. Zhang, and S. He, "Real-time detection network SI-SSD for weak targets in complex traffic scenarios," *Neural Processing Letters*, preprint, 2022.
- [15] A. Vecvanags, K. Aktas, I. Pavlovs et al., "Ungulate detection and species classification from camera trap images using RetinaNet and faster R-CNN," *Entropy*, vol. 24, no. 3, p. 353, 2022.
- [16] J. Redmon and A. Farhadi, "Yolov3: an incremental improvement," 2018, <https://arxiv.org/abs/1804.02767>.
- [17] J. Yuan, X. Ma, G. Han, S. Li, and W. Gong, "Research on lightweight disaster classification based on high-resolution remote sensing images," *Remote Sensing*, vol. 14, no. 11, p. 2577, 2022.
- [18] G. Yang, W. Feng, J. Jin et al., "Face mask recognition system with YOLOV5 based on image recognition," *Farmaco*, in *Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, IEEE, Farmaco, Chengdu, December 2020.
- [19] M. M. Saritas and A. Yasar, "Performance analysis of ANN and Naive Bayes classification algorithm for data classification," *International journal of intelligent systems and applications in engineering*, vol. 7, no. 2, pp. 88–91, 2019.
- [20] D. Song, Y. Tang, and J. Feng, "Aggregating minutia-centred deep convolutional features for fingerprint indexing," *Pattern Recognition*, vol. 88, pp. 397–408, 2019.
- [21] Y. Tang, F. Gao, J. Feng, and Y. Liu, "FingerNet: an unified deep network for fingerprint minutiae extraction," *Farmaco*, in *Proceedings of the 2017 IEEE International Joint Conference on Biometrics (IJCB)*, IEEE, Denver, CO, USA, October 2017.
- [22] J. M. Smith, "Analyzing the mosaic structure of genes," *Journal of Molecular Evolution*, vol. 34, no. 2, pp. 126–129, 1992.
- [23] B. Jin, L. Cruz, and N. Gonçalves, "Deep facial diagnosis: deep transfer learning from face recognition to facial diagnosis," *IEEE Access*, vol. 8, pp. 123649–123661, 2020, 2020.
- [24] Z. Huang, Z. Zhong, L. Sun, and Q. Huo, "Mask R-CNN with pyramid attention network for scene text detection," *Farmaco*, in *Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Farmaco, HI, USA, January 2019.
- [25] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp, "Image segmentation with a bounding box prior," in *Proceedings of the 2009 IEEE 12th international conference on computer vision, F*, IEEE, Kyoto, Japan, September 2009.