IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 6 July 2023, accepted 14 August 2023, date of publication 4 September 2023, date of current version 7 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3311474

RESEARCH ARTICLE

Image Tampering Recognition Algorithm Based on Improved YOLOv5s

ZHEN LIU
Network and Computing Center, Shenyang Institute of Engineering, Shenyang, Liaoning 110136, China
e-mail: Liuzhen@sie.edu.cn

This work was supported in part by the Liaoning Province Education Administration under Grant LJKMZ20220610.

ABSTRACT In order to improve the extraction effect of the features of the tampered region, the model usually relies on a specific feature defined by hand to optimize the model. However, the effect of the model is not ideal when the tampered mode is unknown. In this paper, the Neck layer of the YOLOv5s model is reasonably embedded into the CBAM attention module, so that the model can more accurately capture the features of the tampered region. In addition, the boundary frame loss is optimized by using EIOU loss function, which separately dissolves the impact factors of aspect ratio to calculate the model. The experimental results show that the image tampering recognition algorithm proposed in this paper based on the improved YOLOv5s can not only effectively identify a variety of tampering modes, but also increase the average accuracy of recognition by 1.57% compared with the benchmark algorithm. The recognition speed is higher than that of LSTM and U-Net algorithms, and the recognition speed also reaches 13.89 images per second, maintaining a high level.

INDEX TERMS Image tamper recognition, YOLOv5s, attention module, EIOU loss function.

I. INTRODUCTION

Professor Frid from the Department of Computer Science at State University of New York in the United States and his research team were the first to engage in digital image tamper recognition work in the world. The copy paste tamper recognition algorithm they developed opened the door to digital image recognition. The algorithm first calculated the DCT [1] coefficient, then sorted the image dictionary and recognized the translation between image blocks based on the similarity between image blocks, But this recognition algorithm has a relatively large computational complexity. Later, Professor Farid also proposed using a new method mainly based on principal component analysis theory to significantly reduce the dimensionality of data blocks. This new method mainly uses a large amount of dimensionality reduction coefficient vectors obtained during the quantitative principal component analysis process to directly replace the quantized DCT coefficient vectors, and can use the large amount of dimensionality reduction coefficient vectors obtained as features of image blocks, This method significantly reduces the complexity of the calculation process and can resist certain noise attacks and JPEG compression. Later, researchers proposed a new method to identify tampering in image copying and pasting using the SIFT algorithm. Although this new algorithm has strong robustness to geometric operations on images, its computational time is too long. Bay et al. proposed the SURF algorithm based on the SIFT [2] algorithm, which greatly reduces the complexity of the SIFT algorithm and improves the real-time performance of feature recognition and matching. Mishra et al. proposed a copy and paste recognition method based on accelerated robust feature (SURF) and hierarchical clustering (HAC). Hashmi et al. conducted experiments on SURF algorithm, SURF combined with discrete wavelet transform (DWT), and SURF combined with binary wavelet transform (DYWT) [3], demonstrating that the proposed composite algorithm outperforms individual SURF algorithms in terms of complexity, scale invariance, rotation invariance, and attack combination.

Traditional image tamper detection methods are relatively single in the design of image feature extraction. When the tampering method is unknown, the algorithm effect is not

The associate editor coordinating the review of this manuscript and approving it for publication was Tai-Hoon Kim.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

ideal, and the overall framework of traditional image tamper detection methods is not concise enough. Currently, deep learning based detection methods have been optimized to a certain extent in terms of model and structure. Commonly used algorithms include LSTM [4], RCNN series algorithms, and YOLO series algorithms, However, there is still a lot of room for improvement in the recognition accuracy of the algorithm, especially when identifying small areas of tampering. It is necessary to improve the deep learning algorithm to improve the recognition accuracy of the algorithm. The image tampering recognition algorithm based on the improved YOLOv5s is proposed in this case.

In the digital age, people can easily make false images with exquisite tampering tools, and the Internet is accelerating the spread of false images. From the moral and ethical point of view, image tampering can be divided into benign use and malignant use. Benign uses include entertaining the public, seeking strange pictures or spoofing videos in social activities, etc. The pursuit of visual aesthetics, beautification and retouching in fashion photography and artistic creation, and cool special effects in science fiction movies and TV shows. Malicious use is against ethics. False images with strong deception may sometimes cause incalculable social security problems, such as false propaganda in political news, false pictures in scientific research and academic research, and false identity in personal life. Therefore, in order to maintain social order and public trust, ensure national security and stability, safeguard judicial justice and authority, crack down on illegal and criminal and avoid bad academic research atmosphere, it is an urgent scientific research problem for people from all walks of life to develop reasonable and effective image tampering detection methods to identify the authenticity and reliability of images.

II. EXPERIMENTAL METHOD

A. YOLOv5s

The main idea of YOLO algorithm is to take the entire image as input, extract features through convolutional neural networks, and then directly regression the coordinate values and classification probability of the target. What is significantly different from the two-stage target detection algorithms of the RCNN series is that the YOLO series algorithms provide another approach to identifying targets, omitting the step of predicting candidate box positions and treating the entire problem as a regression problem. By using a single convolutional neural network, end-to-end model training can be completed, significantly improving the speed of object detection algorithms, which far exceeds other real-time detection algorithms.

The basic idea of YOLO algorithm is to directly regress the position coordinates and classification probability of the target suggestion box. Its main feature is that each regressed target is based on the pixel information of the entire image, and will not be misreported in the background, which can better distinguish between the detected object and the background area. Unlike the Faster RCNN algorithm, YOLO algorithm can use global features of the entire image, but the previous versions of YOLO also have lower positioning accuracy than Faster RCNN. The improved version of YOLO has no lower recognition accuracy or speed than the latter.

YOLOv5s [5] is the network model with the smallest network depth and the smallest feature map width among the four models, and its detection precision and speed are excellent. The other three models have been continuously deepened and widened on this basis, and the depth and width of the models can be controlled by code. From the perspective of reducing deployment costs and making the network lightweight, this article selects YOLOv5s, which has the smallest network depth and width, for improvement.

The YOLOv5s network structure can be divided into four parts: input layer, Backbone, Neck, and output layer, as shown in Figure 1.

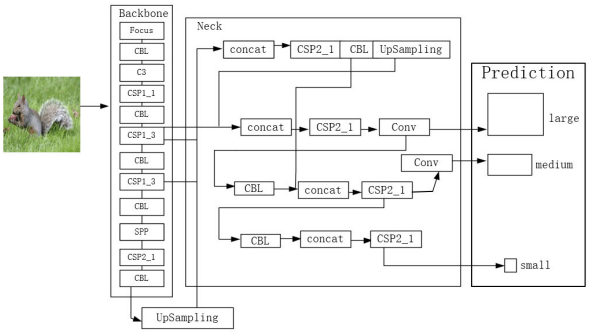

FIGURE 1. YOLOv5s network structure diagram.

The input layer uses mosaic data augmentation technology, which is an ancient data augmentation technique. It refers to using limited data to create as much new data as possible. YOLOv5s model is sent to the network model for training in order in a whole batch. Use hyperparameter to control whether to turn on mosaic data enhancement, randomly take the center point of a generated mosaic data enhanced image, randomly select the index of three other images that need data enhancement, traverse the enhanced image index, and then store the image on the large mosaic image that needs to be generated, perform random rotation, translation, scaling, and cropping data augmentation operations on this image, and finally obtain a mosaic enhanced image. YOLOv5s has two data augmentation forms: Mosaic and Mosaic9, which are aimed at enhancing a batch. This data augmentation method is very helpful for detecting small targets.

The Backbone layer is composed of four modules, namely the Focus module, CBL module, CSP module, and SPP module. Among them, the Focus module adopts slicing operation, first splitting the high-resolution feature maps into multiple low-resolution feature maps, and then performing inter column sampling and stitching operations. The Focus structure is a unique structure of YOLOv5s model, which means that before the image enters the backbone network, sample a value every other pixel of the image. Similar to the adjacent near

down sampling [6] operation, four groups of images can be obtained. Finally, the results obtained will be convolved to obtain the final twice down sampled feature map. The CBL module is composed of three parts: the convolution layer, the batch standardization layer and the activation function layer. The SPP module adopts four pooling cores of different sizes for maximum pooling operation, and then splices the results to obtain the fused features. CSP1_ X and CSP2_ X module draws inspiration from the idea of CSPNet, which is composed of three modules: convolutional layer [7], CBL module, and Res Unit module. The proposal of CSPNet is mainly to enhance the learning ability of CNN, maintain certain accuracy while lightweight deployment, reduce computational costs, and maintain low memory overhead.

The Neck layer adopts a structure that combines FPN and PAN, combining the conventional FPN layer with a bottom-up feature pyramid to fuse the extracted semantic features with positional features. At the same time, the backbone layer is fused with the detection layer to obtain richer feature information for the model.

The output layer outputs prediction boxes at three scales: large, medium, and small. The NMS [8] (non maximum suppression) algorithm is used to remove excess boxes, determine the accurate position of the predicted object, and ultimately complete the entire prediction process.

B. CBAM ATTENTION MODULE

Attention model has been widely used in various fields of deep learning in recent years. Whether in the field of image processing, speech recognition, or natural language processing tasks, we can see the practical application of attention mechanism [9] and get good results. In current object detection technologies, there are already many visual attention models used to focus attention on a certain area in the image, using different weight parameters to adjust the importance of attention information. The important information to be focused on is multiplied by a higher weight value to highlight the importance of this position. For some non critical information, it is filtered by multiplying it by a lower weight value to improve the robustness of the model. For more complex recognition tasks, attention mechanisms can be divided into channel attention mechanisms, spatial attention mechanisms, and mixed attention mechanisms based on their effects on different positions.

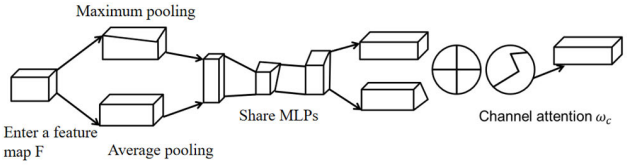

FIGURE 2. Channel attention structure diagram.

The channel attention mechanism focuses more on the relationship between each channel in the feature map, solving the problem of ''what''. It treats each channel as a feature extractor, and its network structure is shown in Figure 2. From the figure, it can be seen that channel attention first performs maximum pooling and average pooling operations on the input feature map, respectively, average pooling is to ensure that each pixel in the feature map can have an impact on the output result, while maximum pooling is to consider the feedback at the position where the maximum response is generated in the feature map during gradient retrieval. The expression of channel attention mechanism can be shown by formula (1-1):

$$\begin{aligned} W_c(F) &= \sigma((\mathrm{MLP}(\mathrm{AvgPool}(F)) + \mathrm{MLP}(\mathrm{MaxPool}(F))) \\ &= \sigma\left(Q_1\left(Q_0\left(F_{ag}^{l}\right)\right) + Q_1\left(Q_0\left(F_{mx}^{l}\right)\right)\right) \end{aligned} \quad (1\text{-}1)$$

wherein,W_c represents the channel attention value, F represents the original feature map, σ is a sigmoid function,MLP is a multi-layer perceptron, F_{ag}^{l} and F_{mx}^{l} are the results of the feature map after average pooling and maximum pooling, respectively. When two one-dimensional vectors are obtained, input the shared MLP (MultilayerPerception) layer, add the output, and finally use the sigmoid function to normalize to obtain the weight value of each channel W_c, The product of channel weight W_c and input feature map F' is the output feature map of channel attention mechanism.

Channel attention mechanism will give a high weight value to the feature channel containing very important Semantic information, which can effectively improve the feature extraction ability of the network model. However, channel attention mechanism only considers channel information, and the size of the feature graph is becoming smaller and smaller with the increase of network layers, leading to the loss of important edge information. As a result, a spatial attention mechanism emerged, which solves the problem of ''where''. The spatial attention network structure is shown in Figure 3.

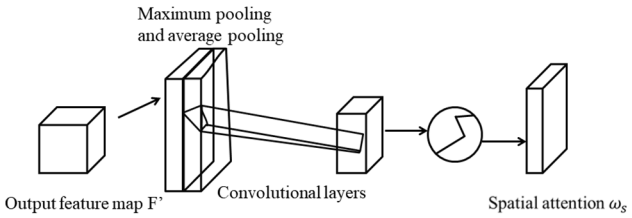

FIGURE 3. Spatial attention structure diagram.

As shown in the figure, the spatial attention mechanism involves performing maximum pooling and average pooling operations on the feature maps along the channel axis. The feature maps obtained from these two pooling operations are concatenated together, and the concatenated feature maps are then convolved. Finally, the sigmoid function is input to obtain the spatial attention weights W_s, W_s and the product of the input feature maps is used as the output feature map of the spatial attention mechanism [10]. The expression of channel

attention mechanism can be shown by formulas (1-2):

$$W_s(F') = \sigma\left(f^{7*7}([AvgPool(F); MaxPool(F)])\right) = \sigma\left(f^{7*7}\left(\left[F^s_{ag} + Q_1\left(Q_0\left(F^s_{mx}\right)\right)\right)\right) \quad (1\text{-}2)$$

Among them, W_s represents the spatial attention value, F is the input feature map, σ is the sigmoid function, F^s_{ag} and F^s_{mx} are the results of the feature map after average pooling and maximum pooling, respectively.

The CBAM attention module is composed of two completely independent sub modules, spatial attention and channel attention, and its network structure is shown in Figure 4.

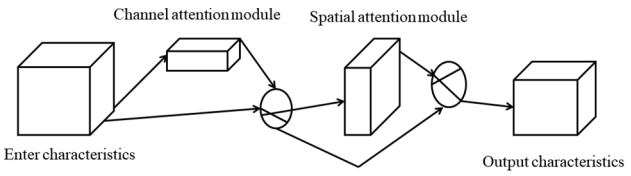

FIGURE 4. CBAM structure diagram.

The overall process of convolutional attention mechanism is to first input the feature map into the channel attention module, obtain the channel weight value, and multiply it with the original input feature. Then, the output of the channel attention is divided into two parts. One part is fed into the spatial attention module, and the obtained spatial attention force weight value is multiplied with the other part to obtain the final output feature map.

Adding the CBAM [11] module to the feature processing Neck structure of the YOLOv5s model, the Neck structure fully integrates the features extracted by the backbone network before sending them to the output layer for prediction. Therefore, the processing ability of the Neck structure on the extracted features directly affects the performance of the algorithm. Adding the CBAM module in front of the CBL module enhances the network's attention to important channels and areas in the feature map, and improves the model's analysis ability for complex scenes, By utilizing limited computing resources, more effective information can be obtained and the detection accuracy of the model can be improved.

C. OPTIMIZATION OF LOSS FUNCTION

The loss function of YOLOv5s network model includes three parts, namely, classification loss, confidence loss and positioning loss. Classification loss is used to distinguish whether the classification of anchor frame and calibration is correct, and confidence loss is used to calculate the confidence of neural network. Among them, the positioning loss was optimized using CIOU loss [12], and the formulas (1-3) are as follows:

$$L_{ciou} = 1 - IOU + \frac{m^2\left(n, n^{gt}\right)}{q} + \alpha v \quad (1\text{-}3)$$

wherein,$m^2\left(n, n^{gt}\right)$ represents the European distance between the central store of the prediction box and the real box, q represents the diagonal distance of the minimum closure area that can contain both the prediction box and the real box,αrepresents the hyperparameter, and v represents the difference in aspect ratio. Compared with the previous loss function, the convergence speed and detection accuracy of CIOU loss have been significantly improved, but the v in the formula reflects the difference in the aspect ratio, rather than the real difference between the width and height and their confidence [13], which sometimes hinders the optimization similarity of the model. To solve this problem, EIOU lose is used as the loss function of the bounding box, as shown in formula (1-4).

$$L_{eiou} = L_{iou} + L_{dis} + L_{asp} = 1 - IOU + \frac{m^2\left(n, n^{gt}\right)}{q^2} + \frac{m^2\left(w, w^{gt}\right)}{q_w^2} + \frac{m^2\left(h, h^{gt}\right)}{q_h^2} \quad (1\text{-}4)$$

wherein, w and h represent width and height respectively, and q is consistent with formula (1-3). The EIOU loss function separates the aspect ratio influence factor to calculate the width and height of the target frame and the anchor frame, which avoids the problem that the width and height of the anchor frame cannot increase or decrease at the same time in the CIOU loss function, and the EIOU loss function has a faster convergence [14] speed, which improves the recognition effect of the target area as a whole.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In order to enable the network to learn more general tampering features, image tamper recognition methods will be experimented and evaluated on three publicly available datasets, namely Columbia, CASIA, and Coverage. Randomly select 3000 images from these three datasets, including images that were tampered with in three ways: copy-paste, stitching, and erasing. The dataset is divided into a 9:1 training set and a testing set, and the images are annotated using the specialized annotation tool LabelImg. This experiment was conducted on the open-source Ubuntu operating system, using the current popular Python deep learning framework to train and test the network model, The graphics card used in this experiment is NVIDIA GeForce GTX1080ti, which uses a small batch processing training method to train sample data. The batch size is set to 4, the number of iterations is 20000, and the learning rate is set to 0.0005. Every 2000 iterations, the learning rate is set to one tenth of the previous time. Figure 5 shows the experimental effect of the algorithm proposed in this chapter. The first column in the figure is the original image, and the second column is the tampered image, The third column shows the detection effect of the algorithm proposed in this chapter. From the figure, it can be seen that the algorithm proposed in this chapter can correctly identify the type of tampering and frame the tampered area.

This article compares several traditional tamper detection methods and deep learning methods, namely ELA [15],

TABLE 1. Comparison of experimental results.

Algorithm	mAP	AUC	F1
ELA	23.53%	42.76%	21.53%
CFA1	26.62%	48.68%	27.62%
NOI1	31.56%	50.21%	17.56%
LSTM	45.29%	58.76%	49.86%
U-Net	47.08%	52.97%	50.31%
YOLOv5s	52.35%	53.77%	64.17%
Improved YOLOv5s	53.92%	54.64%	64.47%

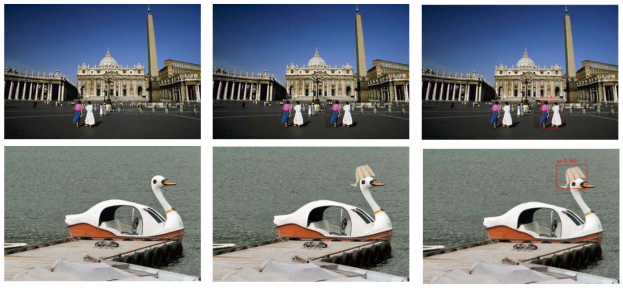

FIGURE 5. Experimental results of the algorithm proposed in this paper.

CFA1 [16], and NOI1 [17]. ELA refers to a tamper recognition method specifically based on error level analysis and judgment, which can distinguish compression differences between tampered and real regions through different JPEG compression quality, and determine whether an image has been tampered with based on this difference. NOI1 uses noise inconsistency to determine whether an image has been tampered with, and the local noise signal of the image is modeled by the high-pass wavelet coefficient. CFA1 is a pixel estimation method based on CFA mode, which uses neighboring pixel values to estimate the filter array mode of the camera. LSTM [4] is an algorithm based on LSTM network to train tampered edges and achieve tamper recognition [18], while U-Net [19] is based on U-Net network to extract features for tamper recognition. The experimental results are shown in Table 1, where the first column represents the algorithm used in the experiment, the second column represents the average accuracy (mAP) of the corresponding algorithm, the third column represents the AUC value of the corresponding algorithm, and the last column represents the F1 value of the algorithm.

From the data in the table, it can be seen that the proposed method has an average recognition accuracy of 30.39%, 27.3%, and 22.36% higher than the three traditional algorithms, respectively. Compared with the other two image tamper recognition methods based on deep learning, the average recognition accuracy is also 8.36% and 6.84% higher respectively, and is 1.57% higher than the average accuracy of the improved YOLOv5s algorithm. On this basis, it can still maintain high AUC and F1 values. Therefore, the method proposed in the paper has a significant improvement compared to other methods, as it can effectively identify the tampered area and locate the location of the tampered area.

IV. CONCLUSION

This paper proposes an image tamper detection algorithm based on improved YOLOv5s. The algorithm integrates CBAM attention module into the Neck layer of the model. The Neck layer fully fuses the features extracted from the backbone network and sends them to the output layer, which enhances the network's attention to important channels and important areas in the feature map, improves the ability to extract features of tampered areas, and uses the EIOU loss function to improve the convergence speed and accuracy of the network, improved the model's ability to analyze complex scenarios. Compared with the traditional tamper detection methods ELA, CFA1, and NOI1, the method proposed in this paper has improved the average accuracy by 30.39%, 27.3%, and 22.36%, respectively. At the same time, the average accuracy of the improved YOLOv5s algorithm has increased by 1.57%.

REFERENCES

[1] M. Tagliasacchi, G. Valenzise, and S. Tubaro, "Hash-based identification of sparse image tampering," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2491–2504, Nov. 2009, doi: 10.1109/TIP.2009.2028251.

[2] P. C. Ng, "SIFT: Predicting amino acid changes that affect protein function," *Nucleic Acids Res.*, vol. 31, no. 13, pp. 3812–3814, Jul. 2003, doi: 10.1093/nar/gkg509.

[3] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, in Lecture Notes in Computer Science, vol. 3951, 2006, pp. 404–417, doi: 10.1007/11744023_32.

[4] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019, doi: 10.1162/neco_a_01199.

[5] C. Jiang, H. Zhang, E. Zhang, Z. Hui, and Y. Yue, "Pedestrian and vehicle target detection algorithm based on the improved YOLOv5s," *J. Yangzhou Univ., Natural Sci.*, vol. 25, no. 6, pp. 45–49, 2022, doi: 10.19411/j.1007-824x.2022.06.008.

[6] F. Cheng, Z. Fu, B. Niu, L. Huang, X. Ji, and Y. Sun, "Fusion of domestic high resolution remote sensing images based on the non-subsampled shearlet transform," *Laser Optoelectron. Prog.*, vol. 59, no. 12, p. 1228001, 2022.

[7] G. Zhang, "Research on remote sensing image scene classification based on deep learning lightweight convolutional neural network," M.S. thesis, Nanjing Univ. Posts Telecommun., 2023, no. 2, doi: 10.27251/d.cnki.gnjdc.2022.000193.

[8] X. Huang, Z. Ge, Z. Jie, and O. Yoshie, "NMS by representative region: Towards crowded pedestrian detection by proposal pairing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 10750–10759, doi: 10.1109/CVPR42600.2020.01076.

[9] H. Zhang, L. Feng, Y. Hao, and Y. Wang, "Ancient mural dynasty identification based on attention mechanism and transfer learning," *Computer*, vol. 43, no. 6, pp. 1826–1832, Mar. 2023, doi: 10.11772/j.issn.1001-9081.2022071008.
[10] Y. Chen, L. Liu, V. Phonevilay, K. Gu, R. Xia, J. Xie, Q. Zhang, and K. Yang, "Image super-resolution reconstruction based on feature map attention mechanism," *Appl. Intell.*, vol. 51, no. 7, pp. 4367–4380, Jul. 2021, doi: 10.1007/S10489-020-02116-1.
[11] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19, doi: 10.1007/978-3-030-01234-2_1.
[12] G. Cai, J.-H. Ciou, Y. Liu, Y. Jiang, and P. S. Lee, "Leaf-inspired multiresponsive MXene-based actuator for programmable smart devices," *Sci. Adv.*, vol. 5, no. 7, Jul. 2019, Art. no. eaaw7956, doi: 10.1126/sciadv.aaw7956.
[13] Z. Lin et al., "Interval state estimation of electricity-gas integrated energy system based on model-data joint driven," *Grid Technol.*, pp. 1–15, Mar. 2023, doi: 10.13335/j.1000-3673.pst.2022.1185.
[14] P. McEnroe, S. Wang, and M. Liyanage, "A survey on the convergence of edge computing and AI for UAVs: Opportunities and challenges," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15435–15459, Sep. 2022, doi: 10.1109/jiot.2022.3176400.
[15] L. Xu, "Image forgery detection based on ELA and CrCbGLCM," M.S. thesis, Huazhong Univ. Sci. Technol., 2020, no. 5, doi: 10.27157/d.cnki.ghzku.2020.005682.
[16] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva, "Image forgery localization via fine-grained analysis of CFA artifacts," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 5, pp. 1566–1577, Oct. 2012, doi: 10.1109/TIFS.2012.2202227.
[17] B. Mahdian and S. Saic, "Using noise inconsistencies for blind image forensics," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1497–1503, Sep. 2009, doi: 10.1016/j.imavis.2009.02.001.
[18] T.-H. Gao, F.-S. Yang, and G.-R. Sheng, "A novel image forensic method based on coefficient-pair histogram in DCT domain," *Optoelectron. Lasers*, vol. 25, no. 11, pp. 2196–2202, 2014, doi: 10.16136/j.joel.2014.11.023.
[19] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021, doi: 10.1109/ACCESS.2021.3086020.

ZHEN LIU was born in 1969. He received the master's degree from Jilin University, in 2010.

He is currently a Senior Laboratory Master with the Shenyang Institute of Engineering. His research interests include school wireless network construction, media resource production, security monitoring, campus one-card deployment, VR technology, and film production. He presided over the completion of the school's wireless network construction, standardized examination room construction, campus security monitoring construction, campus one-card deployment and hardware and software upgrade of the college's network teaching platform, and launched the virtual laboratory platform project, which visualizes abstract and profound learning content by creating a diverse learning environment. He also presided over the construction of national, provincial, and municipal high-quality courses and resource sharing courses in the school, and he has completed more than 400 hours of filming tasks. The film and television features and variety shows he produced as the director have been broadcast on national, provincial, and municipal TV stations. He has rich experience in information construction and teaching resource base construction.