Medicon Engineering Themes
Volume 3 Issue 6 December 2022
Article Type: Review Article
ISSN: 2834-7218

Identification of bird's nest hazard level of transmission line based on improved yolov5 and location constraints

Yang Wu*, Qunsheng Zeng, Peng Li, Wenqi Huang, Lingyu Liang and Jiajie Chen

Digital Grid Research Institute, China Southern Power Grid, Beijing, China

***Corresponding Author:** Yang Wu, Digital Grid Research Institute, China Southern Power Grid, Beijing, China.

Received: November 02, 2022; **Published:** November 19, 2022

Abstract

Bird's nest is a common defect in transmission line, which seriously affects the safe and stable operation of the line. This paper presents a method of bird's nest hazard level identification based on improved yolov5 and location constraints, which solves the problem of bird's nest multiple identification and hazard level classification. We integrate GhostModule and efficient channel attention (ECA) to design a lightweight attention mechanism convolution module (LAMCM). The original yolov5 is improved by using LAMCM and adding a prediction head, which improves the detection ability of small targets and alleviates the negative impact of scale violence. We only identify the bird's nest on the panorama of Unmanned Aerial Vehicle (UAV) patrol, and classify the hazard level of the bird's nest according to the location constraints of the bird's nest and insulator. Experiments on coco dataset and self built transmission line dataset (TL) show that our algorithm is superior to other commonly used algorithms. On the COCO dataset, our algorithm achieved 49.1% AP at a real-time speed of ~79FPS on Tesla V100. On the TL dataset, the recognition effect of our algorithm on towers, insulators and bird nests has improved to varying degrees compared with the original yolov5. In particular, the recall rate of bird nest identification for three hazard levels has increased by more than 3%. The average recall rate of bird nest hazard level identification is 93%, and the average accuracy rate is 93.5%.

Keywords: bird's nest; transmission line; object detection; attention mechanism

Introduction

In recent years, with the rapid development of power grid, overhead high-voltage lines are also increasing. As one of the important components of power system, high-voltage lines shoulder major tasks. In case of any problem, it may cause regional power failure, or endanger the safety of life and property. With the enhancement of people's awareness of environmental protection, the ecological environment is getting better and better, and the number of birds is also increasing. At the same time, there are also some problems. The activities of birds seriously affect the normal operation of transmission lines. The nesting behavior of birds will pollute the insulators in transmission lines, and also cause short circuit or tripping of lines. Therefore, in order to ensure the safe operation of the transmission system, using UAV to patrol the transmission line and realize the accurate detection of bird's nest defects based on image is an urgent problem to be solved in intelligent power operation and maintenance. However, the complex background of patrol images, the change of light intensity, motion blur, target occlusion and scale change all bring great challenges to the detection.

Bird's nest recognition is a typical object detection problem. It needs to solve the location and category of defective targets in the patrol image, that is, to solve the problem of "what and where". Many researchers have carried out image-based bird's nest recognition research, which can be roughly divided into two categories: traditional image detection algorithm and deep learning image detection algorithm. Traditional target detection methods use artificially designed feature descriptors to extract the features of the image, and recognize the targets in the image through sliding window and classification model. The biggest disadvantage of this kind of methods is that the artificially designed feature descriptors have poor universality and poor robustness of the algorithm, which is not suitable for bird's nest detection in complex scenes. By learning a large number of training data, the deep learning target detection algorithm makes the deep convolution neural network have the ability to extract target features and automatically identify and mark bird's nest defects. The application of deep learning technology significantly improves the recognition effect and has good robustness.

The existing bird's nest recognition algorithms have the problem of multiple recognition of the same bird's nest. During patrol inspection, the UAV takes a panoramic image, a tower base image and multiple local images of a tower. The panoramic image is the complete transmission tower photographed by the UAV in the distance, and the local image is the detailed drawing of insulator, connecting hardware and other components photographed by the UAV in the near place. Because there are multiple local images, the scenes between the images may overlap, resulting in a bird's nest may appear in different local images at the same time. Figure 1 shows a local image of a tower at different angles. These images all contain the same bird's nest. Existing algorithms will detect bird's nest defects in all images, which will lead to the problem of multiple identification and repeated recording of defects in the same bird's nest. In addition, it is difficult to see the specific position of the bird's nest on the tower at a glance on the local images, and the position of the bird's nest cannot be accurately described during defect recording.

Figure 1: The same nest photographed from different angles.

The existing identification algorithms do not classify the hazard level of bird's nest. In reality, most towers have birds nesting, and in order to protect birds, humans will specially place some mesh devices for birds to nest. Therefore, some non hazardous nests do not need to be handled in time, and some hazardous nests need to be handled immediately. The bird's nest above the insulator is considered to be a level I hazard, because feces and dripping water can easily lead to the decline of insulator insulation and affect the safe and stable operation of the line. The bird's nest on the cross arm and away from the insulator hanging point is considered to be a level II hazard. The bird's nest in the mesh device is considered to be a level III hazard, also known as artificial bird's nest. Level III hazards basically have no impact on the safe operation of transmission lines. Figure 2 shows bird nests with different hazard levels.

Figure 2: From left to right, there are level I hazards, level II hazards and level III hazards.

In order to solve the problem of bird's nest multiple identification and hazard level division, this paper only focuses on the bird's nest in the panoramic image, and divides the hazard level of bird's nest by the position constraints of bird's nest and insulator string. In the panoramic image, the tower, insulator and other components are intact, and there is only one panoramic image for a tower. Due to the short shooting distance of local image, the shooting of insulator string is often incomplete, and the bird's nest and insulator generally do not exist at the same time in the local image. Therefore, the positional relationship between bird's nest and insulator string cannot be determined in the local image.

As mentioned above, we need to detect the tower, insulator and bird's nest in the panoramic image, which is a typical small target detection problem in the large-resolution image (the proportion of bird's nest in the image is about 1%). In addition, the scale size difference between the three types of targets is obvious, which brings many challenges to the detection. To solve the above problems, referring to the bneck module design in MobileNetv3 [1], we design a lightweight attention mechanism convolution module (LAMCM) integrating GhostModule [2] and ECA [3] to improve the feature extraction ability of the network. The original yolov5 is improved by using LAMCM and adding a prediction head, which improves the detection ability of small targets and alleviates the negative impact of scale violence. We further improve the detection ability of small targets by cut large image into small image and increasing the network input size. Experiments on coco dataset and self built transmission line dataset (TL) show that our algorithm is superior to other commonly used algorithms.

We summarize our contributions as follows:

- We propose to identify the bird's nest hazard level only on the panoramic image, which solves the problem of bird's nest multiple identification and hazard level classification.
- Combining GhostModule and ECA, we designed a lightweight attention mechanism convolution module (LAMCM) to enhance the feature extraction ability of small targets.
- We use LAMCM and add a prediction head to improve the original yolov5, enhance the detection ability of small targets, and alleviate the negative impact of scale violence.

Related Work

Two-stage object detection

Two-stage object detection consists of two steps: candidate regions proposal generation and label classification and location regression were performed on these candidate regions. Faster region-based convolutional neural networks (Faster RCNN) [4] used the region proposal network to find out the candidate regions of possible objects in the image, and then used two head networks to assign the correct category labels to these candidate regions and adjust the position of these candidate regions again. In [5], Dai et al presented region-based fully convolutional networks (R-FCN) for accurate and efficient object detection, R-FCN used position-sensitive score maps to address a dilemma between translation-invariance in image classification and translation-variance in object detection. Cascade RCNN [6] improves the quality of candidate frames generated by Region Proposal Network (RPN) in Faster RCNN, making the final target detection frame positioning more accurate. It consists of a series of detectors, each detector is trained based on positive and negative samples with different Intersection-over-Union (IOU) thresholds. The later the detector, the greater the IOU threshold defining positive and negative samples, the output of the former detector is used as the input of the latter detector, so the detectors are trained stage by stage.

One-stage object detection

Different from the two-stage object detection algorithm, the one-stage object detection completely eliminates proposal generation and directly regress and classify the anchors get the the final bounding boxes. Single shot multibox detector (SSD) [7] combines predictions from multiple feature maps with different resolutions to naturally handle objects of various sizes. In addition, it encapsulates all computation in a single network, making it easy to train. You only look once (YOLO) [8] predicts bounding boxes and class probabilities directly from full images, it is extremely fast. RetinaNet [9] found that the main reason why the accuracy of the one-stage detectors is

lower than that of the two-stage detectors is the extreme foreground-background class imbalance encountered during training, the paper introduces focal loss to address the class imbalance problem by reshaping the standard cross-entropy loss.

Anchor-free object detection

Anchor-free object detection abandons anchors and transforms the object detection problem into key point and scale estimation. CornerNet [10] detect an object bounding box as a pair of keypoints, the top-left corner and the bottom-right corner, using a single convolution neural network. ExtremeNet [11] uses bottom-up approaches to detect four extreme points (top-most, leftmost, bottom-most, right-most) and one center point, and then group the five keypoints into a bounding box if they are geometrically aligned. Centernet [12] uses keypoint estimation to find the center point and regression width and height of the object bounding box. Fovea Box [13] predicts category-sensitive semantic maps for the object existing possibility, and produces category-agnostic bounding box for each position that potentially contains an object. RepPoints [14] represents the targets as a set of sample points useful for both localization and recognition, it learn to automatically arrange themselves in a manner that bounds the spatial extent of an object and indicates semantically significant local areas.

Identification of bird's nest

With the wide application of UAV in transmission line inspection, more and more researchers participate in image-based power grid defect recognition. Wang et al. [15] proposed a solution for identification of bird's nests on transmission lines based on feature identification, which locates the pole tower of transmission line by line segment detector (LSD) line detection, Harris corner detection and morphological closing operation, and detects the bird's nests within the range of pole tower based on their shape and color features, in order to identify the bird's nests accurately. In [16], Li et al. proposed a deep learning-based birds' nests automatic detection framework-region of interest (ROI) mining Faster RCNN. In [17], Lu et al. proposed a novel method of bird's nest images detection based on cascade classifiifier and combination features. In [18], Zhang et al. used RetinaNet model based on deep convolution neural network to automatically detect of bird's nest targets. In [19], Yang et al. proposed a bird's nest location recognition method based on YOLOv4-tiny.

Method

Lightweight Attention Mechanism Convolution Module (LAMCM)

GhostModule is a lightweight channel adjustment module proposed in GhostNet [2], as shown in Figure 3.First, the conventional 1x1 convolution is used to compress the channel number of the input feature map, and then 3x3 depth separable convolution is used to obtain ghost feature map. Finally, concat the identity mapped features and ghost features as the output feature map. GhostModule not only reduces the number of parameters and training time, but also increase the receptive field of the network and enhance the ability of feature extraction because of the introduction of 3×3 depth separable convolution.

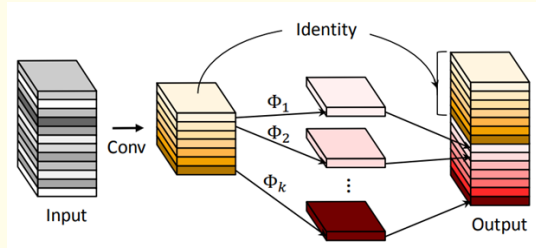

Figure 3: The Ghost module.

Efficient channel attention (ECA) is a channel attention module proposed in ECA-Net [3], as shown in Figure 4. On the one-dimensional feature map obtained after global average pooling (GAP), ECA learns the weights of each channel through a one-dimensional convolution with shared weights. The size of the one-dimensional convolution kernel k represents the cross channel information interaction rate of the module, and k will be dynamically adjusted with the change of the number of channels. ECA solves the problem of information loss caused by the reduction of channel dimension of SE, and greatly reduces the amount of parameters.

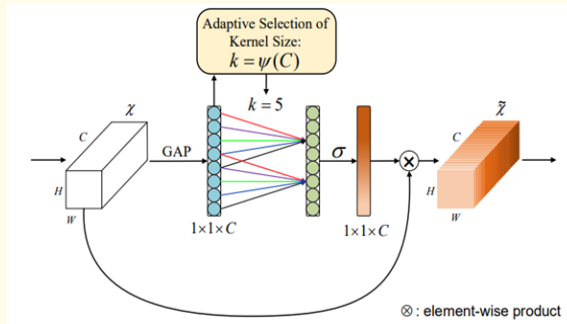

Figure 4: The ECA module.

Referring to the bneck module design in MobileNetv3 [1], we design a lightweight attention mechanism convolution module (LAMCM) by integrating GhostModule and ECA, and its structure is shown in Figure 5.Mobilenetv3 mainly uses 1×1 convolution to expand the number of channels. GhostModule is adopted in LAMCM, while ensuring a small number of network parameters, the combination of 3×3 depth separable convolution and 1×1 convolution can also obtain richer features and larger receptive fields than 1×1 convolution, which is of great benefit to the task of object detection. Compared with the SE module used by MobileNetv3, LAMCM adopts a more lightweight attention module ECA, which further reduces the number of network parameters and improves the performance of the network. LAMCM adds residual connection in the convolution block with stride=1 to solve the problem of gradient disappearance of shallow network during network training.

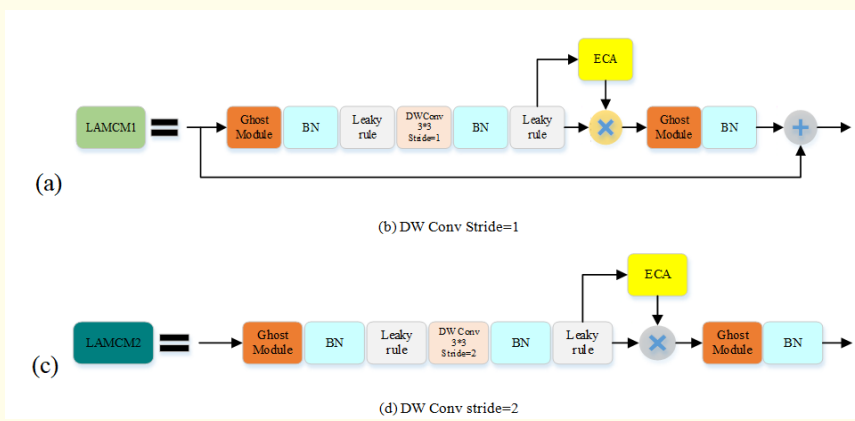

Figure 5: Structure of LAMCM.

Improved YOLOv5

In order to solve the problem of small target detection and large difference in the size of different types of objects under the complex background of transmission panoramic image, we propose an improved YOLOv5 algorithm based on YOLOv5l [20]. The network structure is shown in Figure 6. The improved yolov5 includes three parts: Backbone, Neck and Prediction. The backbone network is composed of four modules: Focus, LAMCM2, CSP1_X, SPP and CSP2_X, which is mainly responsible for the extraction of image features and semantic information. The original Conv + BN + Leaky rule (CBL) in Focus, SPP, CSP1_X and CSP2_X is replaced with lightweight GhostModule + BN + Leaky rule (GBL), and the residual component in CSP1_X is replaced with LAMCM1, which can ensure that the backbone network has smaller network parameters and richer extracted features. Neck part adopts FPN and PAN for multi-scale feature fusion. The embedding of LAMCM1 and LAMCM2 further pays more attention to the channel features with large amount of information, so as to improve the efficiency of feature use. The prediction part adopts four prediction heads, which is one more prediction head than the original YOLOv5l, which can alleviate the negative impact of scale violence, enable the network to accurately predict targets with different scales and improve the detection ability of small targets. We use GhostModule and 1×1 convolution offsets the increase in computing and storage costs caused by adding a prediction header.

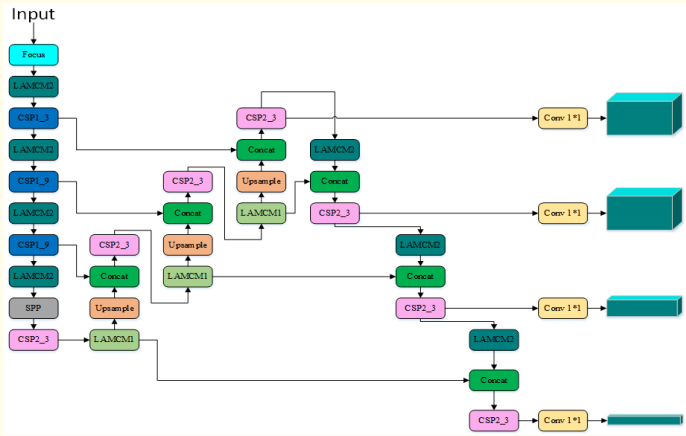

Figure 6: The architecture of the Improved YOLOv5.

Bird's nest hazard level identification

The bird's nest at different positions on the transmission tower has different hazard levels to the line, and the problem that the same bird's nest can be recognized multiple times can only be solved by bird's nest defect recognition in the panoramic image. The bird's nest hazard level identification process is as follows: 1) input the UAV patrol image of the transmission line. 2) The patrol images are 4K or 8K high-resolution large images. The large image are divided into small images. In order to avoid some objects being cut off by segmentation (except towers), two adjacent small images need to be partially overlapped. This step can improve the detection rate of small objects, but it will increase the recognition time of a single image. This step is optional. 3) The improved yolov5 algorithm is used to detect seven kinds of objects: transmission tower, bird's nest, artificial bird's nest, horizontal glass insulator, vertical glass insulator, horizontal composite insulator and vertical composite insulator on large image and small images. There is basically no bird's nest above the horizontal glass insulator and the horizontal composite insulator. The coordinates of the detection objects on the small images are mapped to the high-resolution large image, and then the final detection results is obtained by NMS together with the detection results on the large image. 4) Judge whether there is tower in the detection results. If there is transmission tower, enter the next step. If there is no tower, exit directly. The image without tower is a local image, which avoids the hidden danger identification of bird's nest on the local image. 5) If the overlapping area between the bird's nest and the tower is equal to 0, the bird's nest is considered to

be misidentified, and the misidentified bird's nest is deleted. 6) Expand the left and right frames and upper frames of vertical glass insulator and vertical composite insulator outward by half. 7) Calculate the overlap ratio R between the bird's nest and the extended vertical insulator. 8) The bird's nest is divided into hazard levels. If R is greater than 0, the bird's nest is a level I hazard. If R is equal to 0, the bird's nest is a level II hazard and the artificial bird's nest is a level III hazard.

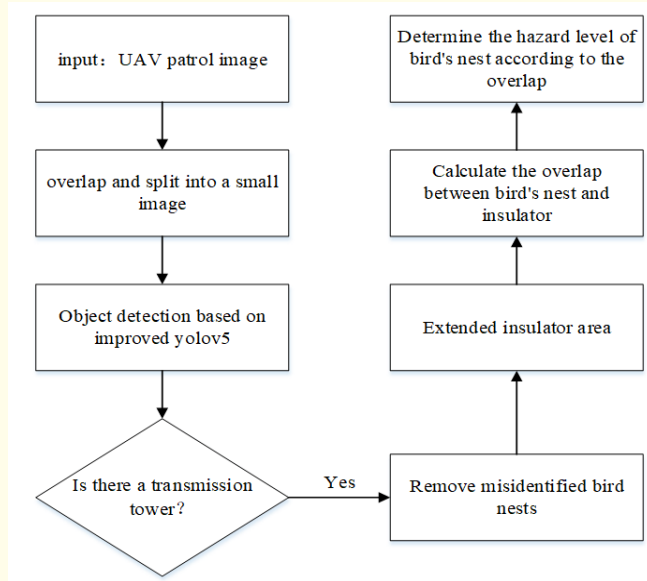

Figure 7: Bird's nest hazard level identification flow chart.

Experiments

We use COCO 2017 dataset and transmission line dataset to evaluate our proposed algorithm, which is the same or slightly improved compared with other common target detection algorithms.

Implementation Details

We implement our algorithm on Pytorch 1.8.0, All of our models use Tesla V100 for training and testing. We train the models for a total of 300 epochs, and the first 5 epochs are used for warm up. We use stochastic gradient descent (SGD) for training, and lr=0.01 as the initial learning rate with the cosine lr schedule. The weight decay is 0.0005 and the SGD momentum is 0.937. The batch size is 128 by default to typical 8-GPU devices. On the coco datasets, the input image size of our model is 640, and the batch size is 128 on 8-GPU devices. On the transmission line UAV patrol datasets, the input image size of our model is 1280, and the batch size is 32 on 8-GPU devices.

Transmission line datasets (TL)

TL dataset is a data set established by ourselves, the pictures in the data set are all from the real images taken by the UAV when patrolling the transmission line. Table 1 lists the number of pictures and labels in the training set and test set. There are 3907 pictures in the training set and 850 pictures in the test set. We use label me tool to label all pictures, the labeling information includes 8 types of labels, namely transmission tower (TT), vertical glass insulator (VGI), horizontal glass insulator (HGI), vertical composite insulator (VCI), horizontal composite insulator (HCI), bird's nest of level I hazard (I-H), bird's nest of level II hazard (II-H) and bird's nest of level III hazard (III-H).

Label Category	Training Set		Test Set	
	Number of Pictures	Number of labels	Number of Pictures	Number of labels
TT	3907	4271	850	883
VGI		10824		697
HGI		8554		351
VCI		13767		1052
HCI		4173		169
I-H		496		182
II-H		2627		416
III-H		1463		375

Table 1: TL dataset.

Comparisons with the state-of-the-art on COCO dataset

On the COCO 2017 dataset, we use FPS, AP_{50} and to evaluate our algorithm and several algorithms commonly used in practical projects. In order to fairly compare with the original yolov5l, our improved yolov5 also adopts three prediction headers, and the input size is 640. It can be seen from table II that the improved yolov5 algorithm has certain advantages in three indicators. Compared with the original yolov5l, our model AP is improved by 0.9 and FPS is improved by 6.6. It can be seen that the improved yolov5 proposed in this paper is effective.

Method	Size	FPS	AP(%)	AP_{50}
Faster-RCNN w/FPN [21]	1000*600	15.6	36.2	59.1
Cascade-RCNN [6]	1312*800	12.4	42.8	62.1
SSD512 [7]	512*512	28.0	28.8	48.5
YOLOv3+ASFF* [22]	608*608	45.5	42.4	63.0
FCOS [23]	1333*800	8.8	43.2	62.8
YOLOv4 [24]	608*608	62.0	43.5	65.7
YOLOv5l [20]	640*640	73.0	48.2	66.9
Improved YOLOv5(three prediction headers)	640*640	79.6	49.1	67.6

Table 2: Comparison of the speed and accuracy of different object detectors on COCO 2017 test-dev.

Comparisons with the state-of-the-art on TL dataset

All images in TL dataset are high-resolution large images. Training or prediction directly with the large images will lead to non convergence of the model and poor detection effect of small targets. During training, we use a 3200×3200 size window to randomly slide around the labeled objects to generate a small image. We can cut more small images for the categories with a small number in the training set to ensure sample balance. During the test, we use a 3200×3200 size window to slide on the large image to generate a small image. There is overlap between adjacent small images, and the overlapping pixels are 640. The model input size is 1280. In order to better detect targets with different scales, we add a prediction head to the model. By judging that the bird's nest does not fall in the tower detection box, we remove the wrongly identified bird's nest and reduce the wrong alarm. We use recall rate and accuracy rate to evaluate the performance of the original yolov5l and improved yolov5, as shown in Table 3. If the IOU of the detection box by the algorithm and the manually labeled box is greater than 0.6, the recognition is correct. For fair comparison, the other settings of the two algorithms are the same except that there is one more prediction header in the improved yolov5. It can be seen from the table that the improved yolov5 algorithm in this paper has advantages in the recognition of all categories. In particular, the recall rate of small targets

such as bird's nest has increased significantly. The recall rate of bird's nest with level I hazard has increased by 4.4%, that of bird's nest with level II hazard has increased by 3.4%, and that of bird's nest with level III hazard has increased by 4.8%.

Label Category	*YOLOv5l*		*Improved YOLOv5*	
	Recall rate	*Accuracy rate*	*Recall rate*	*Accuracy rate*
TT	99.1%	98.5%	99.8%	99.3%
VGI	93.7%	94.3%	94.4%	95.3%
HGI	92.8%	91.1%	93.2%	91.9%
VCI	96.3%	98.5%	97.5%	98.8%
HCI	90.2%	96.9%	90.5%	97.5%
I-H	91.2%	93.3%	95.6%	94.1%
II-H	90.1%	89.7%	93.5%	90.9%
III-H	86.4%	94.7%	91.2%	96.3%

Table 3: Performance comparison between original yolov5l and improved yolov5 on TL dataset.

Figure 8 shows the identification results of some samples on the test set. Because the original image is a high-resolution image, the target is too small to display. The following figure shows enlarged screenshots. Blue is the tower detection box, orange is the insulator detection box, red is the bird's nest detection box with different hazard levels, and the text next to the box describes the defect category. It can be seen from the figure that the algorithm proposed in this paper can accurately locate the area of the tower and reduce the interference of background information. Our algorithm can also identify different types of insulators, which provides a basis for the hazard classification of bird's nest. The shape and size of the bird's nest in the figure are obviously different, and some bird's nests are blurred, but our algorithm can accurately identify the hazard level. Experiments show that the algorithm has strong applicability and high practical value.

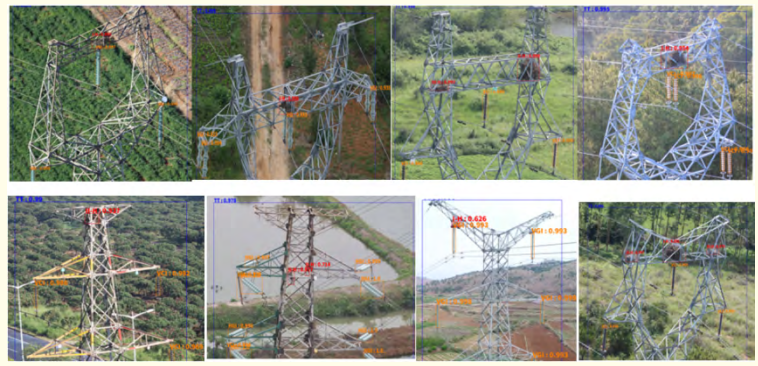

Figure 8: Identification results of some samples on the test set.

Ablation Studies

We analyze importance of each proposed component on TL dataset. We only analyze the overall recall rate and accuracy rate of bird's nest hazard level, because the improvement of accuracy rate and recall rate of other categories is not very obvious, and the index of bird's nest hazard level is the focus of power inspectors. The impact of each component is listed in the table 4. LAMCM represents a lightweight attention mechanism convolution module, APH represents adding a prediction header, and DELETE represents delete misidentification.

Methods	*Recall rate*	*Accuracy rate*
YOLOv5l	88.9%	90.4%
YOLOv5l+LAMCM	91.2%	90.9%
YOLOv5l+LAMCM+APH	93.0%	91.7%
YOLOv5l+LAMCM+APH+DELETE	93.0%	93.5%

Table 4: Ablation study on the total recall and accuracy of bird's nest hazard level on TL dataset.

Conclusion

In this paper, we propose a transmission line bird's nest hazard level identification method based on improved yolov5 and location constraints, which solves the problem of bird's nest multiple identification and hazard level classification. We propose a lightweight attention mechanism convolution module and add a prediction head to improve the original yolov5, which improves the detection ability of small targets and alleviates the negative impact of scale violence. Experiments on COCO dataset and self built transmission line dataset (TL) show that the improved yolov5 proposed in this paper is better than the original yolov5.On the COCO dataset, our algorithm achieved 49.1% APat a real-time speed of ~79FPS on Tesla V100. On the TL dataset, the recognition effect of our algorithm on towers, insulators and bird nests has improved to varying degrees compared with the original yolov5. In particular, the recall rate of bird nest identification for three hazard levels has increased by more than 3%. The average recall rate of bird nest hazard level identification is 93%, and the average accuracy rate is 93.5%.

References

1. Howard A., et al. "Searching for mobilenetv3". Proceedings of the IEEE/CVF International Conference on Computer Vision (2019): 1314-1324.
2. Han K., et al. "Ghostnet: more features from cheap operations". Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020): 1580-1589.
3. Wang Q., et al. "Supplementary material for 'eca-net: efficient channel attention for deep convolutional neural networks". Technical report, Technical report (2020).
4. Ren S., et al. "Faster R-CNN: towards real-time object detection with region proposal networks". Advances in Neural Information Processing Systems 28 (2015).
5. Dai J., et al. "R-FCN: object detection via region-based fully convolutional networks". Advances in Neural Information Processing Systems 29 (2016).
6. Cai Z and Vasconcelos N. "Cascade R-CNN: delving into high quality object detection". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018): 6154-6162.
7. Liu W., et al. "SSD: single shot MultiBox detector". Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9905, pp. 21-37. Springer, Cham (2016).
8. Redmon J., et al. "You only look once: unified, real-time object detection". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016): 779-788.
9. Redmon J and Farhadi A. "Yolov3: an incremental improvement". arXiv preprint arXiv:1804.02767 (2018).
10. Bochkovskiy A, Wang C-Y and Liao H-YM. "Yolov4: optimal speed and accuracy of object detection". arXiv preprint arXiv:2004.10934 (2020).
11. Ge Z., et al. "Yolox: exceeding yolo series in 2021". arXiv preprint arXiv:2107.08430 (2021).
12. Wang C-Y, Bochkovskiy A and Liao H-YM. "Scaled-yolov4: scaling cross stage partial network". Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2021): 13029-13038.
13. Lin T-Y., et al. "Focal loss for dense object detection". Proceedings of the IEEE International Conference on Computer Vision (2017): 2980-2988.

14. Law H and Deng J. "CornerNet: detecting objects as paired keypoints". Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) Computer Vision - ECCV 2018. LNCS, Springer, Cham 11218 (2018): 765-781.
15. Zhou X, Zhuo J and Krahenbuhl P. "Bottom-up object detection by grouping extreme and center points". Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2019): 850-859.
16. Zhou X, Wang D and Krähenbühl P. Objects as points. arXiv preprint arXiv:1904.07850 (2019).
17. Kong T., et al. "Foveabox: beyound anchor-based object detection". IEEE Trans. Image Process 29 (2020): 7389-7398.
18. Yang Z., et al. "Reppoints: point set representation for object detection". Proceedings of the IEEE/CVF International Conference on Computer Vision (2019): 9657-9666.
19. Wang Q. "A solution for identification of bird's nests on transmission lines with UAV patrol". 2016 International Conference on Artificial Intelligence and Engineering Applications, Atlantis Press (2016): 202-206.
20. Ni H, Wang M and Zhao L. "An improved faster R-CNN for defect recognition of key components of transmission line". Math. Biosci. Eng 18.4 (2021): 4679-4695.
21. Li J., et al. "Deep learning-based bird's nest detection on transmission lines using UAV imagery". Appl. Sci 10.18 (2020): 6147.
22. Jianfeng L., et al. "Detection of bird's nest in high power lines in the vicinity of remote campus based on combination features and cascade classifier". IEEE Access 6 (2018): 39063-39071.
23. Hui Z., et al. "Intelligent bird's nest hazard detection of transmission line based on retinanet model". Journal of Physics: Conference Series, IOP Publishing 2005 (2021): 012235.
24. Yang Q., et al. "Lightweight bird's nest location recognition method based on yolov4-tiny". 2021 IEEE International Conference on Electrical Engineering and Mechatronics Technology (ICEEMT), IEEE (2021): 402-405.
25. https://github.com/ultralytics/yolov5.
26. Lin T-Y., et al. "Feature pyramid networks for object detection". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017): 2117-2125.
27. Liu S, Huang D and Wang Y. "Learning spatial fusion for single-shot object detection". arXiv preprint arXiv:1911.09516 (2019).
28. Tian Z., et al. "FCOS: fully convolutional one-stage object detection". Proceedings of the IEEE/CVF International Conference on Computer Vision (2019): 9627-9636.

Volume 3 Issue 6 December 2022