Research on the computer intelligent recognition of electric appliance by Yolo V5 algorithm

Xiaolong Zhang*, Bin Qian, Haitao Si, Liuying Zeng, Hui Wang
Longyou Power Supply Company of State Grid Zhejiang Province, Quzhou, China

*Corresponding author: xiaolongzhang@zj.sgcc.com.cn

Abstract. In the process of power operation, the correct identification of tools can lay a foundation for the detection of violations in power operation. In order to realize the recognition of power instruments, based on the current Yolo V5 algorithm, a detection algorithm for power instruments is proposed by improving Yolo V5 algorithm. Firstly, the model of Yolo V5 algorithm is introduced. Then the establishment of the power tool database and the process of model training are analysed. Finally, the test results are analysed and evaluated. The models generated after training were accelerated by TensorRT and then deployed on Jetson Xavier NX.

Keywords: Yolo V5, Power, Grounding line, Electroscope, Jetson Xavier NX.

1. Introduction
In the process of electric operation, various kinds of instruments are commonly used, and the violations in the process of electric operation often involve some instruments. In order to better supervise the violations of operators, this paper aims to identify the electric instruments in the process of operation, and lays a foundation for the detection of violations. With the continuous optimization and update of power hardware equipment, there are more and more types of power instruments, which brings some difficulties for identification and detection. Traditional machine learning algorithms cannot better distinguish the complicated power instruments. At present, there are few studies on the identification and classification of power instruments, which cannot better promote the analysis of power violations.

At present, deep learning detection technology is widely used in the power industry, among which Yolo series algorithm is one of the important models for engineering application and implementation. In this paper, based on Yolo V5 algorithm, the recognition and classification of power instruments are realized. Yolo V5 algorithm enhances the data of the power tool model, and migrates and learns the data set through preprocessing to expand the database and further ensure the accuracy of the tool recognition.

2. Principles and Methods
Yolo series algorithms have undergone a series of evolution, from the initial Yolo V1 to the current Yolo V5. With the evolution of the algorithm itself, the accuracy and accuracy of the current recognition and detection field have been greatly improved.
2.1. Yolo V5 Network architecture
Current deep learning detection algorithms can be divided into one stage (such as SSD, Yolo, etc.) and two stages (Fast RCNN, Mask RCNN, etc.). Yolo series algorithms belong to one stage, and the main advantage of one stage series algorithms is faster computing speed. It is more popular in the project with low target detection accuracy and more conducive to the landing of the project. Currently, the framework of Yolo series has been gradually developed to Yolo V5 (here in after referred to as V5), which has been further optimized and improved on previous frameworks such as V4 and V3. According to the depth of network and the width of feature map, V5 series can be divided into four models: V5s, V5m, V5l and V5x [1] [2]. Finally, the network model will be deployed on the edge processing equipment Jetson Xavier NX. Since the target object is an electric tool, which is not a small object, considering the detection rate and accuracy, the smallest network V5s model can be adopted for the test, and V5s is based on the Pytorch network framework. It can meet the speed and accuracy required by target detection in this paper and is suitable for deployment in edge processing equipment. The network structure of V5s is shown in Figure 1 below.

The features of Yolo V5 are as follows [3] [4] [5]:
(1) The matching strategy of positive sample Anchor in the neighborhood is considered, and the positive sample is increased;
(2) Models with different complexity can be obtained by flexible configuration parameters;
(3) Improve the overall performance through some built-in super parameter optimization strategies;
(4) Like Yolo V4, Mosaic was used to enhance the detection performance of small objects;
(5) Using the latest version of Pytorch for mixing accuracy and distributed training;
(6) Warm up + COS LR learning rate strategy, without weight attenuation for bias;
(7) Adopt the gradient accumulation strategy commonly used in Yolo series, increase batch size, and perform special initialization of bias in the output head part; The class balanced sampling strategy is adopted.
(8) Multi-scale training, but very rough writing, directly to the data loader output batch image bilinear interpolation;
(9) Support on NX format export;
(10) The EMA strategy of moving average of model weight index is adopted.

2.2. Loss function
In the Yolo V5 model, the cross-entropy loss function is used to calculate the score loss, which is represented by L_GIOU. The formula is shown in equation (1) [6], which can be understood in combination with Figure 2 for ease of understanding.

![Figure 1. Network structure of Yolo V5](image-url)
A represents the prediction box, B represents the real box, and C represents the intersection (overlapping region) of A and B.

\[
\begin{align*}
C &= A \cap B \\
IOU &= \frac{|A \cap B|}{|A \cup B|} \\
GIOU &= IOU - \frac{|C/(A \cup B)|}{|C|} \\
L_{GIOU} &= 1 - GIOU = 1 - (IOU - \frac{|C/(A \cup B)|}{|C|})
\end{align*}
\] (1)

Figure 2. Schematic diagram of IOU

In equation (1) above, IOU is the cross-parallel ratio, which can be used to determine positive and negative samples and reflect the detection effect of prediction box and real box. GIOU, as a measure of distance, has scale invariance.

2.3. The activation functions

In Yolo V5 model, Sigmoid is used as the activation function. Sigmoid function is a common S function [3]. The function expression is shown in equation (2) and Figure 3 below.

\[
\sigma(z) = \frac{1}{1+e^{-z}}
\] (2)

Figure 3. Sigmoid function
Sigmoid function is in a smooth state on the whole, with negative infinity tending to 0 and positive infinity tending to 1, so the value of is (0,1), and the function is symmetric about 0. Combined with the above functions, Sigmoid function is insensitive to input beyond a certain range, and data is not easy to diverge during transmission, which provides a good environment for the convergence of Yolo V5 model.

2.4. Model training and edge deployment

The training and development environment of the data model was ubuntu18.04 and Pytorch framework. Two NVIDIA GeForce GTX 2080Ti were used in the test process for pre-processing and training. Finally, the trained model framework was deployed and accelerated on the edge processing equipment. The model is deployed on NVIDIA's edge processor, Jetson Xavier NX (described below as NX), as shown in Figure 4 below, which brings supercomputer performance to the edge. It includes a powerful compact Jetson Xavier NX module for AI edge devices. It benefits from new cloud-native support that accelerates the NVIDIA software stack to 10W, which is more than 10 times the performance of its widely adopted predecessor, the Jetson TX2 [7]. With precise multi-mode AI reasoning capabilities, energy-efficient, small-size solutions can be developed and tested, opening the door to new breakthrough products. Developers can now leverage the benefits of cloud native support to transform the experience of AI software development and deployment to edge devices.

The Jetson Xavier NX offers 14 TOPS at 10 watts and 21 TOPS at 15 watts, making it ideal for systems that are limited in size and power. With 384 CUDA cores, 48 Tensor Core and 2 NVDLA engines, it can run multiple modern neural networks in parallel and process high-resolution data from multiple sensors simultaneously [7].

![Figure 4. Jetson Xavier NX](image)

3. Test results and analysis

3.1. Data set construction and training

3.1.1. Data set construction. Longyou image training data from quzhou electric power company, for the power operation personnel violation detection basis, training data are derived from the process, the database mainly includes the operation personnel use grounding line, electroscope, security fence, such as three conditions, including the correct use and is not the correct use of two kinds of circumstances, such as database image number of a total of 2059 images; The images are divided into safety fence setting, illegal crossing of safety fence, correct use of grounding wire, illegal use of grounding wire, correct use of electroscope, illegal use of electroscope, etc., as shown in Figure 5 below. After finishing the data set, labelimg was used to label the images of objects including security fence, ground wire, electroscope, etc., and the labeling generated.xml file. Since.txt file was used in the training process of Yolo V5 short hair model, it was necessary to convert the.xml file.
3.1.2. Training images. For safety fence, ground wire, electroscope, especially the grounding line and electroscope, with the optimization of the equipment, the types of ground wire and electroscope, bring certain difficulty for the test, according to the problem, on the one hand need to amplify and instrument database, on the other hand, need to find out the same characteristics of all kinds of ground wire and electroscope, can really improve the accuracy of instrument detection.

In the process of image training based on Yolo V5, image data was enhanced, and then transfer learning was used to expand the database. The proportion of training set and verification set was divided according to 8:2. The data set construction and training flow chart are shown in Figure 6 below.

![Flowchart of data set construction and training](image-url)
3.2. Experimental results and analysis

Model training input image size is 640*640, IOU threshold is 0.5, initial learning rate is 0.01, weight attenuation rate is 0.001, iteration times is 40, batch processing size is 8. Image training results are shown in Figure 7 below. The detection accuracy of tape, electroscope and ground wire is 99.5%, 99.7% and 99.6% respectively.

![Figure 7. The training results](image)

Based on the weight of the above image training, the test results are shown in Figure 8 below, from which it can be seen that the ground wire, electroscope, warning band and so on can be correctly detected.
Therefore, Yolo V5 algorithm can be used as an early detection algorithm for power violation analysis. The model is lightweight and easy to deploy on the edge processing facility Jetson Xavier NX.

In the early stage of the deployment of the trained model, in order to improve the real-time performance of the algorithm, the model acceleration tool TensorRT developed by NVIDIA is used. TensorRT mainly inferences the model in the use process. The working principle of TensorRT is to reconstruct the network structure during the reasoning process. The Yolo V5 model used in this paper is based on the Pytorch framework. During the reconstruction process, TensorRT combines some processes that can be combined and calculated, such as the integration of loss function, offset and convolution layer, and the same operation layers at the same level can be combined to accelerate the network model.

4. Conclusion
In summary, this paper uses the Yolo V5 algorithm model to detect power instruments, and the detection accuracy of tape, electroscope and grounding wire is 99.5%, 99.7% and 99.6% respectively. The model is deployed on Jetson Xavier NX equipment, which lays a foundation for the detection of violations in power operations.

Acknowledgments
This work was financially supported by the project research on Power operation and maintenance Behavior Control Technology Based on Edge Computing and Visual Cognition(5211QZ2000U4).

References
[1] Wen H, Dai F, Yuan Y. A Study of YOLO Algorithm for Target Detection [J]. 2021.
[2] Liu K, Tang H, He S, et al. Performance Validation of Yolo Variants for Object Detection[C]/Proceedings of the 2021 International Conference on Bioinformatics and Intelligent Computing. 2021: 239 - 243.
[3] Yang G, Feng W, Jin J, et al. Face Mask Recognition System with YOLOV5 Based on Image Recognition [C]. 2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, 2020.
[4] Ruan Jiyang. Design and Implementation of Target Detection Algorithm based on YOLO [D]. Beijing University of Posts and Telecommunications, 2019.
[5] Kuznetsova A, Maleva T, Soloviev V. Detecting apples in orchards using YOLOv3 and YOLOv5 in general and close-up images[C]/International Symposium on Neural Networks. Springer, Cham, 2020: 233 - 243.
[6] Rezatofighi, Tsoi N, JY Gwak, et al. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.(GIOU).
[7] https://developer.nvidia.com/embedded/jetson-xavier-nx.