energies

MDPI

Article

A Lightweight Electric Meter Recognition Model for Power Inspection Robots

Shuangshuang Song, Hongsai Tian and Feng Zhao *

School of Information Science and Electrical Engineering, Shandong Jiaotong University, Jinan 250357, China; 22208012@stu.sdjtu.edu.cn (S.S.); 23208028@stu.sdjtu.edu.cn (H.T.)
* Correspondence: zhaofeng@sdjtu.edu.cn

Abstract: Power inspection robots are essential for ensuring safe and optimal operation of power systems. However, during the operation of the power inspection robot, constraints imposed by computational and storage resources slow down the detection speed of the power system, failing to meet real-time monitoring requirements. To address these issues, this study proposes a lightweight electric meter recognition model for power inspection robots based on YOLOv5. The aim is to ensure efficient operation of the model on embedded devices, achieve real-time meter recognition, and enhance the practicality of the inspection robot. In the proposed model, GhostNet, a lightweight network, is employed as the YOLOv5 backbone feature extraction module, thus improving the model's computational efficiency. In addition, the Wise-IoU (WIoU) loss function is used to improve the localization accuracy of the electric meter recognition model. Moreover, the GSConv module was introduced in the neck network for further model lightweighting. The experimental results demonstrated that the proposed model achieves a recognition accuracy of 98.8%, a recall rate of 98.8%, and a frame rate of 416.67 frames per second, while reducing computational volume by 25% compared to the YOLOv5 model. Furthermore, through case studies and comparisons, we illustrated the effectiveness and superiority of the proposed approach.

Keywords: power inspection robots; deep learning; electric meter recognition; GhostNet; loss function

check for updates

Citation: Song, S.; Tian, H.; Zhao, F. A Lightweight Electric Meter Recognition Model for Power Inspection Robots. *Energies* **2024**, *17*, 4731. https://doi.org/10.3390/en17184731

Academic Editors: Ahmed F. Zobaa and Paulo Santos

Received: 11 July 2024
Revised: 11 September 2024
Accepted: 18 September 2024
Published: 23 September 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A power inspection robot is an automated device used to inspect, detect, and maintain power equipment, thus enhancing safety and reliability. It is a novel power inspection method capable of performing equipment inspection and maintenance in challenging environments [1]. The electric power inspection robot relies on advanced navigation technology and an environment sensing system, which can realize the function of autonomous path planning. In the planning process, it can analyze environmental information in real-time and flexibly adjust the travel path, thus significantly improving work efficiency. In the face of obstacles or unexpected conditions, the robot can quickly respond to effectively avoid obstacles or adjust the response strategy to ensure the smooth progress of the inspection task. This system not only improves the efficiency of inspection work, but also provides a strong guarantee for the safe and stable operation of the power system [2]. Electric meter recognition technology is essential for power inspection robots as it helps ensure power system safety and efficient energy management [3]. Accurate electric meter recognition helps ensure the stability of the power system and increases energy consumption efficiency [4,5]. However, some challenges are encountered in electricity meter recognition when using power inspection robots. For example, electric meter images may be reflective and dark due to environmental changes. In addition, electric meters are typically located in complex backgrounds with various interfering objects, increasing the complexity of recognition. These characteristics greatly affect the efficiency and accuracy of electric meter recognition.

With advances in deep learning (DL) technology, DL-based electric meter recognition algorithms have received increased research interest in recent years. For example, Huang

et al. proposed an electric meter recognition algorithm based on the multiscale convolutional neural network and long-short-term memory (LSTM) network [6]. This method extracts features by using the multiscale convolutional neural network (CNN) and then models and recognizes sequences by using the LSTM network. This method exhibited good performance on various standard datasets. Gu et al. proposed a deep convolutional neural network (deep CNN)-based power meter recognition algorithm [7]. They used a deep CNN to extract features and a SoftMax classifier for classification and recognition, achieving a recognition rate of over 90% on standard datasets. Li et al. employed ResNet, a deep CNN architecture, for electric meter recognition [8]; they trained it to extract meaningful features from electric meter images by fine-tuning and combining cross-entropy loss, ternary loss, and other loss functions. ResNet addresses the problem of residual connectivity by incorporating a gradient vanishing problem in deep network training. It can learn higher-level feature representations from photos, thereby improving the electric meter recognition ability [9]. Wang et al. proposed a power meter recognition approach based on a CNN and support vector machine (SVM) [10]. They used a CNN to extract features and an SVM classifier for classification and recognition, obtaining good recognition accuracy when used on a power inspection robot. Yang et al. proposed an electric meter recognition method based on the attention mechanism. The attention mechanism adaptively selects feature regions in the image linked to meters, thereby enhancing the electric meter recognition accuracy [11]. The attention module dynamically learns the attention weights for each location in the image, resulting in improved accuracy. Mutis et al. proposed an electric meter recognition method based on DL and target detection [12,13]. They used the YOLOv3 target detection algorithm to localize and recognize electric meters and then used a CNN for classification and recognition, achieving good recognition results when applied to a power inspection robot. Zhao Hui et al. [14] used the improved YOLOv3 algorithm for monitoring and recognizing the indicator instruments in the substation environment, which can meet the actual needs of the substation. The recognition results were 73.7% and 45.8% faster than Faster R-CNN (faster region-convolutional neural network) and original YOLOv3 network, respectively. Tang Peng et al. [15] proposed base mask-region convolutional neural network (Mask-RCNN) for automatic recognition of digital meter readings in offshore booster stations, where the original images of different types of digital meters are made into a dataset, and deep learning algorithms are utilized for training, and then recognition and analysis are performed; for the digital electricity meter, the meter reading area is discriminative to colored light, especially green light. Anis A et al. [16] proposed a YCbCr image meter reading area extraction method based on Cb and Cr values. The Canny edge operator is then utilized for edge detection of the extracted region. In this edge image, the edge gaps of individual digits are filled by morphological operations to find the edge image of individually connected digits. These individual digits are segmented from the edge image by vertical projection and the individually segmented digits are filled and refined to detect the shape of the digits. Horizontal and vertical binary (HVB) pattern features are extracted from these segmented digit shapes. In 2019, Shuo H et al. [17] proposed a DL-based meter reading method for different types of meters. Digital area recognition and meter type recognition by deep learning and image enhancement can eliminate the effects of different appearance of digital screens, different shooting angles, and environmental interference, and improve the adaptability of the method. Experiments show that the method can effectively recognize the digits of different types of meters; Zhang Z et al. [18] proposed analyzing the attributes of the connectivity domain in the image results to remove the interference information in the image data, to determine the category of the digits, and to determine whether or not it contains a decimal point. Experiments show that the detector combining the connectivity domain analysis method can effectively solve the problem of mutual interference between categories in neural networks, thus improving the detection performance.

Although the aforementioned DL-based electric meter recognition algorithms yield promising results, they encounter challenges in real-world scenarios, such as reflections,

dimness, and obstructed dials of electric meters. Addressing these challenges is crucial. Furthermore, improving the recognition accuracy and minimizing the amount of processing required are important research areas. In order to cope with these challenges and further improve the performance and efficiency of the instrument recognition model, this study will optimize the model architecture, reduce the model complexity, and improve the lightweight nature of the instrument recognition model to meet the efficient real-time requirements of power inspection. Specifically, this study aims to achieve the following:

(1) Improve recognition accuracy: improve the accuracy of meter recognition by the power inspection robot by improving the model architecture to ensure accurate recognition in various complex environments;

(2) Make the model more lightweight: optimize the model structure, reduce the model complexity, and reduce the amount of model computation while ensuring the recognition accuracy, so as to ensure that the meter recognition model can run stably on resource-constrained mobile devices;

(3) Enhance real-time performance: by optimizing the algorithm, the inference speed of the model is significantly improved to ensure that the electric power inspection robot can efficiently carry out meter recognition, thus improving the efficiency of electric power inspection.

2. Methods

2.1. YOLOv5 Method

YOLOv5 is a DL-based model used for target detection. It employs a series of convolutional layers, activation functions, and pooling operations to detect targets with improved performance and speed [19]. The model consists of a backbone feature extraction network and a detection head. To extract high-level features from images, the backbone feature extraction network employs pretrained CNN-based models such as ResNet and EfficientNet. The detection head includes convolutional and fully connected layers that generate bounding boxes and predict the target class [20,21]. The model employs a single-stage detection approach, partitioning the entire image into multiple grids. In each grid, a specific region of the target is identified. The target detection task is thus transformed into a regression problem for estimating bounding boxes and class probabilities of targets within each grid [22]. YOLOv5 has a simpler architecture and faster detection speed compared to the standard two-stage technique, thus making it suitable for real-time applications such as power inspection robots [23]. To detect targets of varying sizes, the model employs feature maps of different scales and feature fusion techniques to improve target detection accuracy, thereby enabling power inspection involving electric meters of various sizes. Furthermore, the model can adaptively process images with different resolutions, enabling power inspection robots to handle images of different sizes. Therefore, in this paper, we selected YOLOv5 as the benchmark model to study electric meter recognition algorithms for use in power inspection robots.

2.2. Improvement of YOLOv5 Network Architecture Design

To further improve the model's performance, we incorporated the GhostNet module as the backbone feature extraction module. GhostNet is a lightweight module that reduces the number of parameters and computational complexity while maintaining high feature characterization capability, thereby improving the efficiency and inference speed of the model. It replaces traditional convolutional layers with Ghost Bottleneck blocks [24], each containing low-dimensional and high-dimensional convolutional operations. In addition, we employed WIoU as the loss function instead of the standard CIoU loss function. WIoU is an enhanced intersection and concurrency ratio (IoU) approach that considers weight information when calculating the loss function, enabling a more accurate evaluation of target detection performance. The adoption of WIoU improves model optimization, making it more accurate for electricity meter recognition in power detection robots. Finally, we introduce a new lightweight convolutional method—GSConv in the model neck network, which uses a combination of standard convolution (SC), depth separable convolution (DSC),

and blending to make the output of the convolutional computation as close as possible to the output of SC, reducing the computational cost and alleviating the model complexity. The schematic of the electric meter recognition network structure based on the improved YOLOv5 is shown in Figure 1.

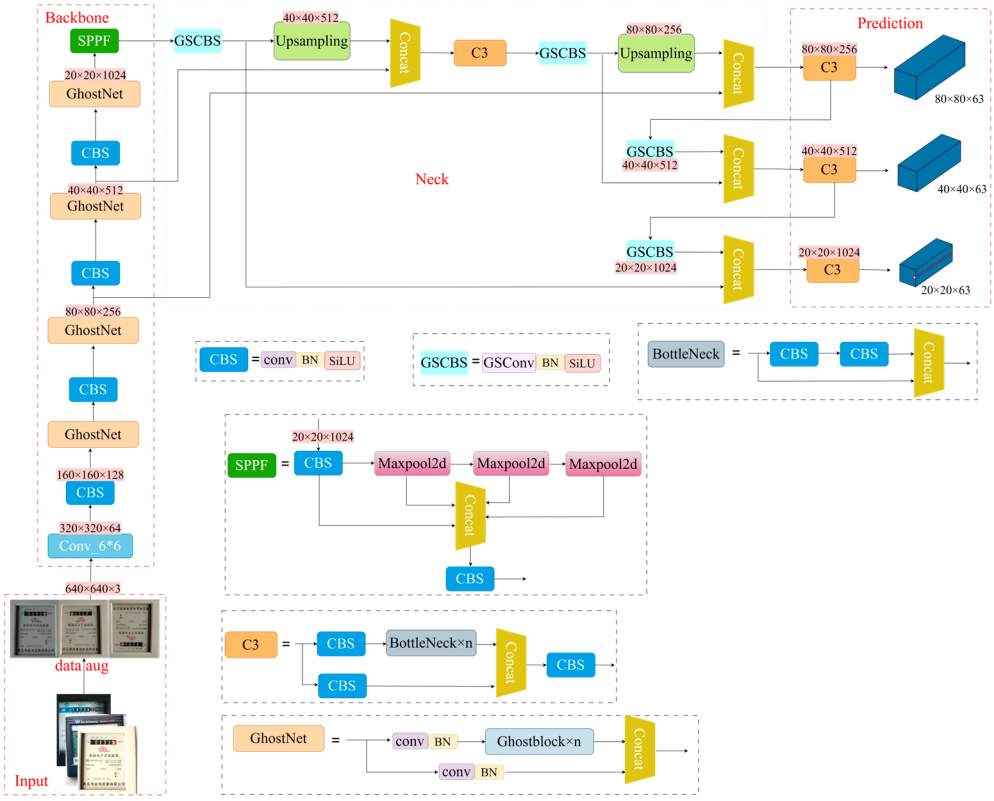


Figure 1. Structure of the electric meter recognition model.

2.2.1. Improvement of the Backbone Network

To reduce the number of parameters in the model, we integrated the Ghost module into the electric power inspection robot for recognizing electric meters. The structure of the Ghost module network is shown in Figure 2. Compared to the traditional convolutional module, the Ghost module reduces the number of model parameters while maintaining a high accuracy rate.

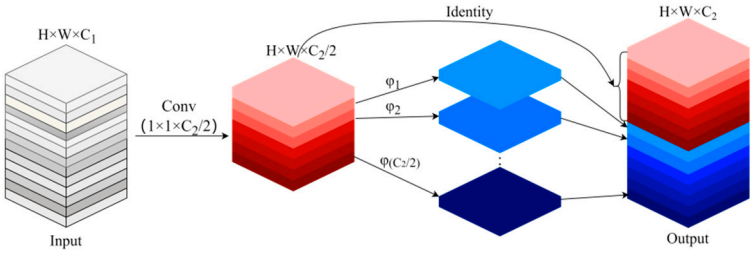
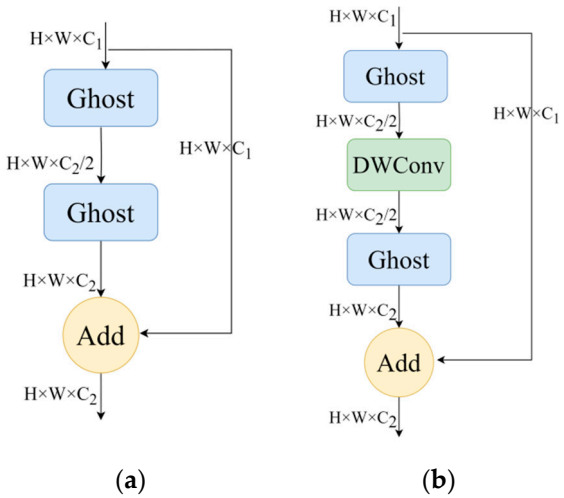


Figure 2. Network structure of the Ghost module.

We achieved a more lightweight design by replacing the original model's backbone feature extraction module with GhostNet. As shown in Figure 3, the GhostNet bottleneck network structure comprises a master branch and a cheap branch. The master branch executes the original convolutional calculations, whereas the cheap branch performs lightweight convolutional computations with fewer output channels [25].



(**a**) (**b**)

Figure 3. Structure of the GhostNet bottleneck network. (**a**) Ghost bottleneck with stride = 1; (**b**) Ghost bottleneck with stride = 2.

As shown in Equation (1), the input data initially undergoes convolutional calculations in the master branch. Subsequently, the cheap branch applies depth-separable convolution to the output of the main branch to generate additional output channels (Equation (2)). Finally, the outputs of the master branch and cheap branch are combined to produce the final output of the Ghost module (Equation (3)):

$$Y1 = X * F1 \times 1 \times C_2/2 \tag{1}$$

$$Y2 = Y1 * F_{dp} \tag{2}$$

$$Y = Concat(Y1, Y2) \tag{3}$$

where $X \in R^{H \times W \times C_1}$ is the input feature, $*$ denotes convolutional operation, $F1 \times 1 \times C_2/2$ is point-wise convolution, $Y1 \in R^{H \times W \times C_2/2}$ is the output feature, F_{dp} is deep-wise convolution, and $Y \in R^{H \times W \times C_2}$ is the output feature, H represents the height of the feature map, W represents the width of the feature map, C_1 represents the number of feature channels of the input image, and C_2 represents the number of feature channels of the output image.

2.2.2. Loss Function Improvement

The loss function is a key component of the target detection model, directly affecting its performance and accuracy. For target detection, the standard YOLOv5 loss function employs complete intersection over union (CIoU) [26,27]. The CIoU loss function is derived as shown in Equations (4) and (5):

$$CIoU = IoU - [\frac{\rho^2(\mathrm{b}, \mathrm{b}^{\mathrm{gt}})}{\mathrm{c}^2} + \alpha v] \tag{4}$$

$$L_{CIou} = 1 - CIoU \tag{5}$$

where c is the diagonal length of the smallest outer bounding box of the real box and the predicted box, (b, b^{gt}) is the distance between the real box and the predicted box's center point, and α is the weight function as stated in Equation (6):

$$\alpha = \frac{v}{(1 - IoU) + v} \tag{6}$$

where v is utilized to measure the similarity of the aspect ratio between the real frame and the predicted frame, as shown in Equation (7):

$$v = \frac{4}{\pi^2}(\arctan\frac{\omega^{gt}}{h^{gt}} - \arctan\frac{\omega}{h}) \tag{7}$$

The CIoU loss function incorporates various concepts, such as centroid distance and width-to-height ratio difference, making it more complex to compute and resulting in a difficult model to train. To address these challenges, we employed WIoU as the loss function in the improved network. WIoU, a bounding box regression loss, includes a dynamic nonmonotonic mechanism and an optimal gradient gain assignment strategy to reduce large or harmful gradients that occur in extreme samples [28]. Its calculation formula is presented in Equation (8):

$$L_{WIoUv3} = r \times L_{WIoUv1} \tag{8}$$

$$L_{WIoUv1} = R_{WIoU} \times L_{IoU} \tag{9}$$

$$R_{WIoU} = e^{\frac{(x-x_{gt})^2+(y-y_{gt})}{(W_g{}^2+H_g{}^2)*}} \tag{10}$$

where L_{WIoUv1} is the W_{IoUv1} with a two-layer attention mechanism obtained by constructing the distance attention based on the distance metric and is computed as shown in Equation (9). R_{WIoU} is its penalization term, which better attenuates the penalization of the geometric metric in the case of a high degree of overlap between the anchor frame and the target frame; it is defined as shown in Equation (10). (x, y) are the coordinates of the centroid of the prediction frame, (x_{gt}, y_{gt}) are the coordinates of the centroid of the prediction frame, and H_g and W_g are the length and width of the minimum outsourcing frame, respectively. The superscript * denotes the separation of H_g and W_g from the computational graph and effectively removes the obstacles to convergence. IoU is used to measure the degree of overlap between the predicted frame and the real frame in the target detection task and is defined as shown in Equation (11).

$$LIoU = 1 - IoU \tag{11}$$

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{12}$$

$$\beta = \frac{L^*_{IoU}}{\overline{L_{IoU}}} \tag{13}$$

Here, r is the gradient gain and is defined as shown in Equation (12), where α and δ are two hyperparameters in the WIoU loss function that can be adjusted to obtain the optimal solution according to different models and datasets, with default values of 1.9 and 3. β is the outlier used to describe the quality of the anchor frame; a small outlier implies that the anchor frames are of high quality, and at this point, the anchor frames receive a small gradient gain to return the bounding box to focus on the anchor frames of ordinary quality. Assigning a small gradient gain to anchor frames with large outliers prevents low-quality examples from generating large harmful gradients (Equation (13)).

The WIoU loss function utilizes a gradient gain allocation strategy combined with a dynamic nonmonotonic focus mechanism to guide the model's attention toward samples of average quality [29], thereby improving the network's generalization capabilities and overall performance.

2.2.3. Improvement of the Neck Network

In order to reduce the computation of the model and improve the detection speed while maintaining the detection accuracy of the model, we introduce the GSConv module into the meter recognition model. The structure of the GSConv module network is shown

in Figure 4. Compared to the SC, GSConv reduces the number of parameters of the model, improves the computational efficiency of the model, and makes the model lighter while maintaining higher performance.

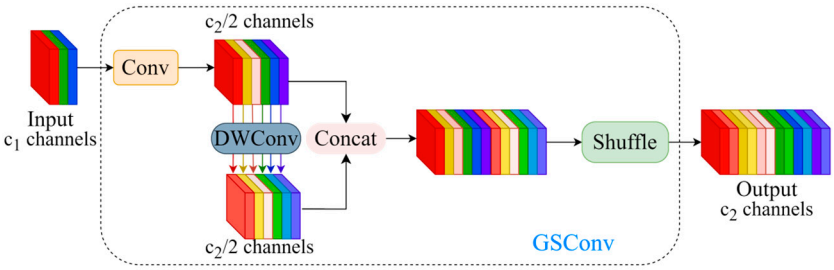


Figure 4. The structure of the GSConv module.

The "Conv" box consists of three layers: a convolutional-2D layer, a batch normalization-2D layer, and an activation layer. The "DWConv" marked in blue here means the DSC operation. DSC operations can significantly reduce parametric and floating-point operations, but DSC separates the channel information of the input image during the computation process, which leads to the feature extraction and fusion ability of DSC being much lower than that of SC. The GSConv module combines SC and DSC, on the basis of which the shuffle operation is introduced to permeate the information generated by SC into every part of the information generated by DSC, which facilitates the exchange of information between the different channels and contributes to the improvement of the model's expressive and generalization abilities [30].

3. Experimental Methodology

3.1. Experimental Setup and Implementation

The experiments were conducted on a server equipped with a PhysX NVIDIA GeForce RTX 4090 GPU. The algorithms were implemented using the PyTorch DL framework, and CUDA was employed to increase the speed and efficiency of training. We prepared two datasets for training three models as follows: one for recognizing numeric dial types, and another for electric meter values. The digital dial type dataset contains approximately 4000 images of electric meters with LCD dials and two types of mechanical dials, labeled as tableone, tabletwo, and tablethree, respectively. Each image contains an electric meter and a dial type. The electric meter value dataset includes nearly 8000 images, labeled with numbers 0 to 9 and decimal places, for a total of 11 categories.

We employed data augmentation techniques such as random rotation, grayscale conversion, mirror transformation, and luminance transformation to increase the diversity and richness of the datasets. This helps prevent overfitting and improves the model's generalization ability. In addition, expanding the dataset reduces the model's training difficulty and enhances training efficiency by providing richer and more diverse samples for the model to learn from, thereby allowing it to recognize and locate the target more accurately.

To ensure fair training, validation, and evaluation, we split the dataset into training, validation, and test sets in the ratio of 8:1:1. The training set was used to train the model, the validation set was used for tuning and early stopping, and the test set was used to evaluate the model's performance. In this study, we combined two models for detecting digital dial types and electric meter values to achieve electric meter value recognition. The digital dial-type model was used to classify the gauges and recognize instrument values based on this classification. This end-to-end process allows direct input of electric meter images and outputs the recognized electric meter values. We trained the model by using the SGD optimizer with an initial learning rate of 0.01, a learning rate decay factor of 0.01, and 100 training epochs. To prevent overfitting, we employed the early stopping technique, wherein training was terminated if the validation loss did not decrease for five consecutive epochs. The graph of train loss and precision changes during training is shown in Figure 5.

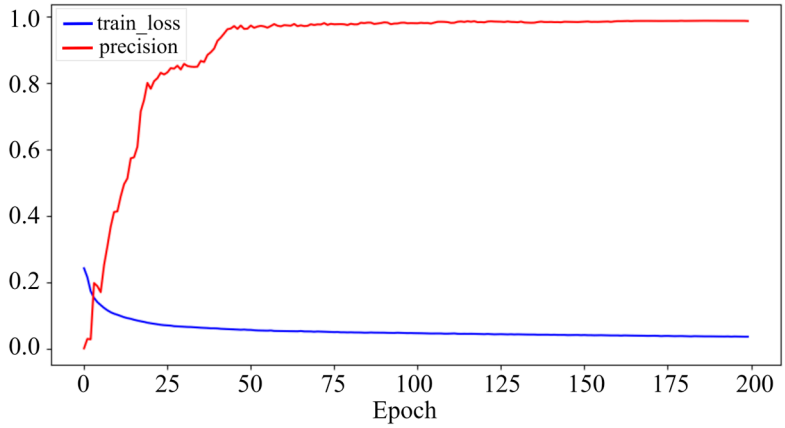


Figure 5. Plot of train loss and precision changes.

3.2. Measurement Indicators

In this study, Precision, Recall, FPR, TNR, F1-Score, mean average precision (mAP) @0.5:0.95, frames per second (fps), and GFlops were employed as measurement indicators to evaluate the accuracy of the improved electric meter recognition model. The accuracy rate is defined as the ratio of accurately predicted positive samples to the total number of expected positive samples and is calculated as follows:

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

Recall is defined as the ratio of accurately predicted positive samples to the total number of actual positive samples and is calculated as follows:

$$Recall = \frac{TP}{TP + FN} \tag{15}$$

FPR (false positive rate) is an important measure of the model's misreporting ability, which indicates the proportion of all samples that are actually negative cases that the model incorrectly predicts as positive cases. The higher this percentage is, the more errors the model makes in distinguishing between negative and positive examples. FPR is defined as shown in Equation (16). TNR (true negative rate) is a measure of the model's ability to correctly identify negative samples. It indicates the percentage of all samples that are actually negative cases that the model is able to correctly predict as negative cases. The higher the TNR, the higher the accuracy of the model in identifying negative cases. The definition of TNR is shown in Equation (17):

$$FPR = \frac{FP}{FP + TN} \tag{16}$$

$$TNR = \frac{TN}{FP + TN} \tag{17}$$

where TP denotes the number of positive samples that were correctly identified, FP denotes the number of negative samples that were incorrectly identified as positive samples, FN denotes the number of positive samples that were incorrectly identified as negative samples, and TN denotes the number of negative samples that were correctly identified as negative samples.

Generally speaking, precision and recall are negatively correlated. In order to balance the impact of precision and recall, F1-Score is introduced as a comprehensive index. F1-

Score is the reconciled average of accuracy and recall, and a larger F1-Score indicates a higher quality of the model.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{18}$$

mAP represents the average precision of the model and is calculated as follows:

$$AP = \int P(R)dR \tag{19}$$

$$mAP = \frac{\sum_{i=1}^{N} APi}{N} \tag{20}$$

When IoU is set as 0.5, mAP@0.5 represents the average AP across all categories. The model's fps is the number of images processed per second, and the processing time for each image includes network inference time as well as non-maximum suppression (NMS) processing time. The computational complexity of the model is measured in GFlops, which is the number of floating-point operations in billions of operations per second.

4. Results and Analysis

4.1. Experimental Results

In the experiments, the models were evaluated using metrics such as precision, recall, and GFlops, and the proposed models were validated using the five-fold cross-validation method to ensure the robustness and generalization of the results, as shown in Table 1. The experimental results show that the optimized algorithm exceeds 98% in terms of precision, recall, and F1 score on the test set. The average FPR of the model is lower than 1%, while the average TNR of the model is higher than 99%, which verifies that the proposed model has excellent performance and stability in distinguishing positive and negative samples. In addition, compared with the original model, the proposed model in this paper reduces the computation by 25% while maintaining accuracy, making it very suitable for deployment in mobile devices such as power inspection robots.

Table 1. Five-fold cross validation performance evaluation table.

Fold	Precision	Recall	F1	mAP@0.5	FPR	TNR
1	0.989	0.99	0.989	0.991	0.006	0.994
2	0.995	0.993	0.994	0.994	0.006	0.994
3	0.991	0.989	0.99	0.993	0.006	0.994
4	0.99	0.987	0.988	0.99	0.007	0.993
5	0.994	0.995	0.994	0.995	0.006	0.994
Average	0.992	0.991	0.991	0.993	0.006	0.994

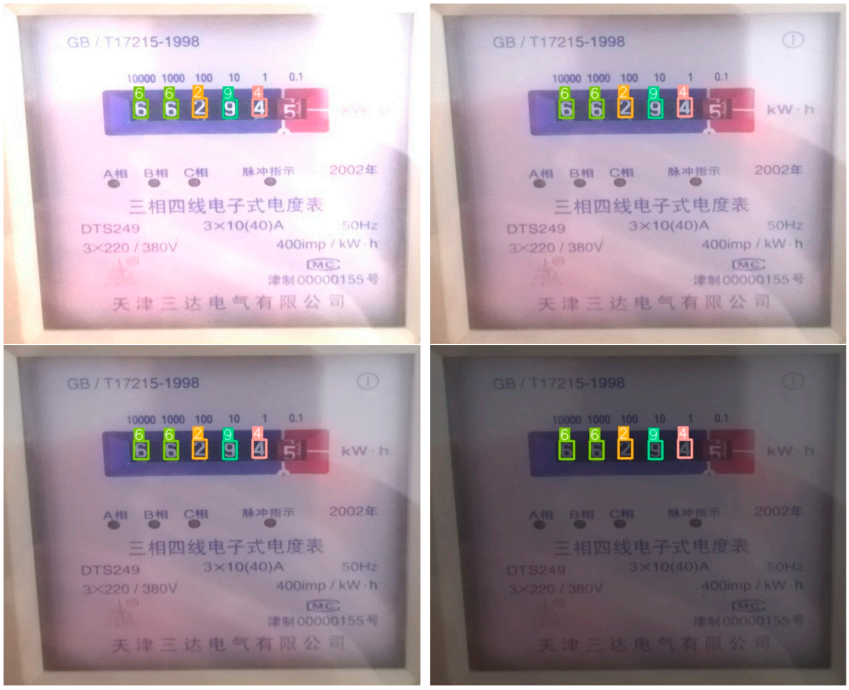
Figures 6–10 show the results of meter recognition under different lighting conditions and different backgrounds as well as different angles, and the experimental results show that the proposed model can accurately identify the meter images mentioned in the above scenarios. In addition, the experimental results were verified by five-fold crossover experiments, as shown in Table 1, which indicate that the proposed meter recognition model has better robustness.



Figure 6. Recognition results for an electric meter with complex background features.



Figure 7. Recognition results for an electric meter with reflective features.



(**a**)



(**b**)

Figure 8. Meter recognition results under different lighting conditions: (**a**) recognition results of mechanical meters under different light conditions; (**b**) recognition results of liquid crystal type meters under different light conditions.



Figure 9. Recognition results of meter images at different shooting angles.



(**a**)

(**b**)

Figure 10. *Cont.*



(**c**)

Figure 10. Recognition results for different types of electric meters: (**a**) Recognition results for Tableone-type dial meters; (**b**) recognition results for Tabletwo-type dial meters; (**c**) recognition results for Tablethree-type dial meters.

4.2. Comparison Experiment

In order to comprehensively assess the performance superiority of the improved network, this study compared it exhaustively with four mainstream target detection networks, and the results are shown in the Table 2 by objectively analyzing the key indexes of each model. Comparison results show that compared with the Faster R-CNN network, the improved network proposed in this study achieves a significant improvement in accuracy, and at the same time has more obvious performance advantages in terms of model complexity and inference speed. Compared with YOLOv5s, the improved network is comparable in terms of accuracy, but the improved network in this study achieves significant optimization in terms of inference speed and model complexity. Moreover, the improved network is slightly higher than the YOLOv7 [31] and YOLOv9 [32] networks in terms of accuracy, and the improved network is more ahead compared to the YOLOv7 and YOLOv9 networks in terms of inference speed and model complexity. In summary, the improved network in this study not only maintains the high accuracy rate of the model, but also reduces the model complexity, greatly improves the processing speed, and meets the stringent real-time requirements of the meter recognition task. Therefore, the improved network has high practical value in the field of meter recognition; the bolded text in the table shows the best results under each indicator.

Table 2. Comparison of performance metrics of different algorithms.

Methods	GFlops	Precision	Recall	F1	mAP@0.5	FPS
Faster R-CNN	183	0.941	0.958	0.949	0. 953	38.46
YOLOv5s	15.8	0.986	0.989	0.987	0.992	135.14
YOLOv7	103.3	0. 978	0.965	0.971	0.984	100
YOLOv9	239	0.982	0.978	0.979	0.985	56.5
Ours	**11.7**	**0.988**	**0.988**	**0.988**	**0.991**	**416.67**

4.3. Comparative Analysis

To demonstrate the superior performance of the optimized model, we conducted comparative experiments on the electric meter reading recognition dataset. First, the original YOLOv5 model was used for training and testing on the meter reading recognition dataset; YOLOv5 served as the baseline model. Next, GhostNet was used as the feature extraction module in the model backbone to reduce the model parameters and computation; next, the loss function was optimized to more accurately measure the difference between the model prediction results and the real labels. Moreover, the GSConv module was used instead of SC in the model neck to further improve the model lightness. Lastly, we validated

the effect of different WIoU versions and the effects on model performance when the two hyperparameters α and δ have different values, and the experimental results are shown in Table 3, the bolded text in the table shows the best results under each indicator.

Table 3. Comparison of model performance for different values of hyperparameters in WIoU.

WIoU Versions	Precision	Recall	F1	mAP@0.5	FPS
WIoUv2 (r = 2)	0.984	0.98	0.982	0. 989	**454.55**
WIoUv3 (α = 2.5, δ = 2)	0.984	0.981	0.982	0.988	400
WIoUv3 (α = 1.9, δ = 3)	0. 982	0.984	0.983	0.989	344.82
WIoUv3 (α = 1.6, δ = 4)	0.986	0.982	0.984	0.985	434.78
WIoUv3 (α = 1.4, δ = 5)	**0.988**	**0.988**	**0.988**	**0.991**	416.67

By analyzing and comparing, when using WIoUv3 and when α = 1.4, δ = 5, the model maintains a higher recognition performance and at the same time, the detection speed is also greatly improved, as shown in Figure 11.

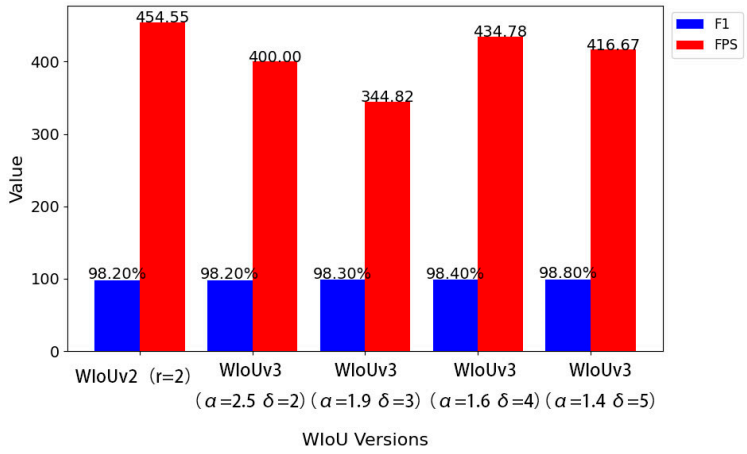


Figure 11. Comparison of F1 and FPS for different values of hyperparameters in WIoU.

The experimental results demonstrated the superior performance of the optimized model for electric meter recognition in power inspection robots. A comparison of the performance of different models on different metrics is presented in Table 4, the bolded text in the table shows the best results under each indicator.

Table 4. Comparison of model performance after adding each module.

Methods	GFlops	Precision	Recall	F1	mAP@0.5	FPS	FPR	TNR
YOLOv5	15.8	0.986	**0.989**	0.987	**0.992**	135.14	0.009	0.991
YOLOv5_GhostNet	12.3	0.987	0.982	0.984	0.991	128.21	0.009	0.991
YOLOv5_GhostNet_WIoU	12.3	0.988	0.979	0.983	0.99	163.93	0.000	1.000
Yolov5-WIoU-GhostNet-GSConv **(proposed method)**	**11.7**	**0.988**	0.988	**0.988**	0.991	**416.67**	0.009	0.991

As can be observed from Table 4, integrating GhostNet as the feature extraction block reduced the model's computational complexity by 20% compared to the original model and improved its computation time. In addition, implementing the WIoU loss function enhanced the recognition performance of the algorithm, achieving an accuracy and recall of 98.8% and 97.9%, respectively, with a computational speed of 163.93 frames per second, making the model more efficient for practical applications. Finally, the GSConv module was introduced to maintain high accuracy while further reducing model complexity and significantly improving computational speed, reaching 416.67 frames per second.

4.4. Experimental Validation in the Power Inspection Robots

The hardware components of the power inspection robots mainly consist of a chassis, motors, processor, and two cameras. The chassis is based on a four-wheel differential drive design with encoders, equipped with the ATmega2560 microcontroller produced by Microcore Technology Co., Ltd. (Beijing, China), combined with the motor drivers, ultrasonic sensors, infrared sensors, angular acceleration sensors, electronic compasses, and Bluetooth communication modules, to achieve various functions such as remote control, autonomous movement, infrared tracking, and obstacle avoidance. Additionally, the chassis design has good expandability and can be equipped with various peripheral modules, such as WiFi video modules, robotic arm modules, GPS positioning modules, and IoT sensor modules, further expanding the functionality and application scope of the intelligent vehicle. The power inspection robots are programmed using the Python programming language to build a highly integrated software platform. This platform includes functions such as data collection, data model construction, autonomous recognition of bends, and unmanned driving verification. Through a distributed structured software design framework, the power inspection robot achieves efficient integration of hardware driver modules, enabling software and hardware to work together to realize various functions of the power inspection robots.

To verify the effectiveness of the proposed algorithm, the trained model for recognizing instruments was transplanted to the power inspection robots. In the verification process, the power inspection robot relies on high-resolution cameras to obtain instrument image information and uses deep learning and visual processing technologies to achieve autonomous operation and instrument recognition functions for the electric power inspection robot. The scene of instrument recognition in the electric power inspection robot is shown in Figure 12.



Figure 12. Electric meter recognition experiment of power inspection robot.

Through experimental validation of the instrument recognition algorithm for electric power inspection robots, the outstanding performance of the proposed algorithm was confirmed. This algorithm can accurately and efficiently recognize various types of instruments while achieving a good balance between recall and precision, effectively avoiding issues of missed detections and false alarms. Through repeated testing, the model demonstrates stable recognition performance under different lighting conditions and complex backgrounds, fully validating the reliability and stability of the proposed instrument recognition algorithm. This experimentally proves the superiority of the algorithm in practical applications. The experimental test results are shown in Figure 13.



Figure 13. Experimental results of electric meter recognition for power inspection robot.

5. Discussion

Electric meter recognition is crucial for power inspection robots as it helps ensure power system safety and efficient energy management. However, environmental changes can lead to reflective or dark electric meter images, making recognition in complex backgrounds with interfering objects challenging. These characteristics significantly affect recognition efficiency and accuracy. Although DL-based electric meter recognition algorithms have yielded promising results, numerous problems are encountered during practical implementation. Problems such as reflections, dimness, and obstructed dials can affect the recognition performance of the algorithm.

To improve computational efficiency, we incorporated GhostNet, a lightweight network, as the YOLOv5 backbone feature extraction module. In addition, we used WIoU as the loss function to enhance the localization accuracy. The effects of different versions of the WIoU loss and different values of the two hyperparameters α and δ on the model performance were compared.

The experimental results demonstrate that the proposed electric meter recognition model achieved an accuracy of 98.8% and a recall rate of 98.8%. It also increased the frame rate to 416.67 frames per second and reduced the computational volume by 25% compared to the YOLOv5 model. Table 4 summarizes these improvements, showing the superior performance of the proposed method over other methods.

In future work, we will enhance the model's accuracy and speed by integrating novel multi-attention mechanisms into the DL model to improve its real-time performance.

6. Conclusions

In this paper, we presented a DL-based electric meter recognition model based on YOLOv5 for power inspection robots. This model addresses the challenges encountered when using conventional YOLOv5 in electric meter recognition. The main conclusions of this paper are as follows:

(1) The lightweight Ghost module was introduced into the YOLOv5 backbone network, which greatly reduced the number of model parameters while enhancing the recognition accuracy and speed;

(2) The WIoU loss function was used for bounding box regression, which improved the stability and localization accuracy of the electric meter recognition model;

(3) Introducing the GSConv module in the YOLOv5 neck network reduced the computational cost of the model and improved its detection speed;
(4) Compared to related YOLOv5 methods, the proposed method exhibited better electric meter recognition performance.

The experimental results demonstrated that the proposed method offers faster recognition speed and higher recognition accuracy. Furthermore, the proposed model exhibits high stability and achieves outstanding performance in electric meter recognition.

Author Contributions: Conceptualization, S.S. and F.Z.; Formal analysis, S.S.; Funding acquisition, F.Z.; Methodology, S.S.; Project administration, F.Z.; Resources, F.Z.; Software, S.S. and H.T.; Validation, S.S. and H.T.; Writing—original draft, S.S.; Writing—review and editing, F.Z. and H.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Shandong Provincial Department of Transportation Science and Technology Plan Project (2023B78-06).

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Katrasnik, J.; Pernus, F. A survey of mobile robots for distribution power line inspection. *IEEE Trans. Power Deliv.* **2009**, *25*, 485–493. [CrossRef]
2. Caban, J.; Nieoczym, A.; Dudziak, A.; Krajka, T.; Stopková, M. The planning process of transport tasks for autonomous vans—Case study. *Appl. Sci.* **2022**, *12*, 2993. [CrossRef]
3. Dong, L.; Chen, N.; Liang, J.; Li, T.; Yan, Z. A review of indoor-orbital electrical inspection robots in substations. *Ind. Robot. Int. J. Robot. Res. Appl.* **2022**, *50*, 337–352. [CrossRef]
4. Alahakoon, D.; Yu, X. Smart electricity meter data intelligence for future energy systems: A survey. *IEEE Trans. Ind. Inform.* **2015**, *12*, 425–436. [CrossRef]
5. Depuru, S.S.S.R.; Wang, L.; Devabhaktuni, V. Smart meters for power grid—Challenges, issues, advantages and status. In Proceedings of the 2011 IEEE/PES Power Systems Conference and Exposition, Phoenix, AZ, USA, 20–23 March 2011.
6. Huang, C.J.; Kuo, P.H. Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting. *IEEE Access* **2019**, *7*, 74822–74834. [CrossRef]
7. Sun, Y.; Zhang, W.; Gu, H.; Liu, C.; Hong, S.; Xu, W.; Gui, G. Convolutional neural network based models for improving super-resolution imaging. *IEEE Access* **2019**, *7*, 43042–43051. [CrossRef]
8. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6999–7019. [CrossRef]
9. Mahdianpari, M.; Salehi, B.; Rezaee, M.; Mohammadimanesh, F.; Zhang, Y. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sens.* **2018**, *10*, 1119. [CrossRef]
10. Wang, F.; Li, K.; Wang, X.; Jiang, L.; Ren, J.; Mi, Z.; Catalão, J.P. A distributed PV system capacity estimation approach based on support vector machine with customer net load curve features. *Energies* **2018**, *11*, 1750. [CrossRef]
11. Yang, M.; Li, X.; Liu, Y. Sequence to point learning based on an attention neural network for nonintrusive load decomposition. *Electronics* **2021**, *10*, 1657. [CrossRef]
12. Mutis, I.; Ambekar, A.; Joshi, V. Real-time space occupancy sensing and human motion analysis using deep learning for indoor air quality control. *Autom. Constr.* **2020**, *116*, 103237. [CrossRef]
13. Zhang, H.; Chen, Q.; Lei, L. A YOLOv3-Based Industrial Instrument Classification and Reading Recognition Method. *Mob. Inf. Syst.* **2022**, *2022*, 7817309. [CrossRef]
14. Zhao, H.; Jiang, L.; Wang, H.; Yue, Y. Inspection of pointer meters based on machine vision. *Sci. Technol. Eng.* **2021**, *21*, 14665–14672.
15. Tang, P.; Liu, Y.; Wei, H.; Dong, X.; Yan, G.; Zhang, Y.; Yuan, Y.; Wang, Z.; Fan, Y.; Ma, P. Automatic recognition algorithm of digital instrument reading in offshore booster station based on Mask-RCNN. *Infrared Laser Eng.* **2021**, *50*, 156–163.
16. Anis, A.; Khaliluzzaman, M.; Yakub, M.; Chakraborty, N.; Deb, K. Digital electric meter reading recognition based on horizontal and vertical binary pattern. In Proceedings of the 2017 3rd International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 7 December 2017; pp. 1–6.
17. Shuo, H.; Yu, X.; Lv, D.; Li, S.; Peng, Y. Digital recognition of electric meter with deep learning. In Proceedings of the 2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), Changsha, China, 1–3 November 2019; pp. 600–607.
18. Zhang, Z.; Hua, Z.; Tang, Y.; Zhang, Y.; Lu, W.; Dai, C. Recognition method of digital meter readings in substation based on connected domain analysis algorithm. *Actuators* **2021**, *10*, 170. [CrossRef]

19. Xie, F.; Lin, B.; Liu, Y. Research on the coordinate attention mechanism fuse in a YOLOv5 deep learning detector for the SAR ship detection task. *Sensors* **2022**, *22*, 3370. [CrossRef]
20. Hasan, M.M.; Chopin, J.P.; Laga, H.; Miklavcic, S.J. Detection and analysis of wheat spikes using convolutional neural networks. *Plant Methods* **2018**, *14*, 1–13. [CrossRef]
21. Carion, N.; Massa, F.; Synnaeve, G.; Usunier, N.; Kirillov, A.; Zagoruyko, S. End-to-end object detection with transformers. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020.
22. Tian, Y.; Yang, G.; Wang, Z.; Wang, H.; Li, E.; Liang, Z. Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput. Electron. Agric.* **2019**, *157*, 417–426. [CrossRef]
23. Liu, M.; Li, Z.; Li, Y.; Liu, Y. A fast and accurate method of power line intelligent inspection based on edge computing. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–12. [CrossRef]
24. Tang, J.; Liu, S.; Zhao, D.; Tang, L.; Zou, W.; Zheng, B. Pcb-yolo: An improved detection algorithm of pcb surface defects based on yolov5. *Sustainability* **2023**, *15*, 5963. [CrossRef]
25. Tang, Y.; Han, K.; Guo, J.; Xu, C.; Xu, C.; Wang, Y. GhostNetv2: Enhance cheap operation with long-range attention. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 9969–9982.
26. Li, Y.; Zhang, J.; Hu, Y.; Zhao, Y.; Cao, Y. Real-time Safety Helmet-wearing Detection Based on Improved YOLOv5. *Comput. Syst. Sci. Eng* **2022**, *43*, 1219–1230. [CrossRef]
27. Chen, Z.; Wu, R.; Lin, Y.; Li, C.; Chen, S.; Yuan, Z.; Zou, X. Plant disease recognition model based on improved YOLOv5. *Agronomy* **2022**, *12*, 365. [CrossRef]
28. Zhou, Y.D.; Zhang, C.Y.; Qiao, Y.H.; Chen, Z.W. Improved Yolov8 for Lightweight Pen Face Keypoints Detection. *J. Hangzhou Dianzi Univ.* **2024**, *44*, 51–64.
29. Saydirasulovich, S.N.; Mukhiddinov, M.; Djuraev, O.; Abdusalomov, A.; Cho, Y.I. An improved wildfire smoke detection based on YOLOv8 and UAV images. *Sensors* **2023**, *23*, 8374. [CrossRef]
30. Zhao, R.; Ma, H.; Xiao, Q.; Shi, G. DIAFNet: A Dynamic Interactive Adaptive Fusion Network Based on Enhanced Differential Features. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 1–5. [CrossRef]
31. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 7464–7475.
32. Wang, C.Y.; Yeh, I.H.; Liao, H.Y.M. Yolov9: Learning what you want to learn using programmable gradient information. *arXiv* **2024**, arXiv:2402.13616.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.