

Article

A Method for Matching Information of Substation Secondary Screen Cabinet Terminal Block Based on Artificial Intelligence

Weiguo Cao ¹, Zhong Chen ^{1,*}, Congying Wu ² and Tiecheng Li ³¹ School of Electrical Engineering, Southeast University, Nanjing 210096, China; caoweiguo_seu@163.com² State Grid Economic and Technological Research Institute Co., Ltd., Beijing 100005, China; wucongying@chinasperi.sgcc.com.cn³ Power Science and Research Institute of State Grid Hebei Power Co., Wuhan 430024, China; ltc8086@163.com

* Correspondence: chenzhong_seu@163.com

Abstract: The matching of schematic diagrams and physical information of terminal blocks in substation secondary screen cabinets plays a crucial role in the operation and maintenance of substations. To enhance the automation level of this task and reduce labor costs, a method for identifying and matching information of terminal blocks in substation secondary screen cabinets based on artificial intelligence is investigated in this paper. Initially, multi-layer object detection networks, tailored to the characteristics of both the schematic diagrams and the physical entities in substation secondary screen cabinets, are designed for the precise extraction of information. Subsequently, network topologies for both the schematic and physical systems are established using the Neo4j database, which allows for the digital storage of information in the substation secondary screen cabinet systems. Finally, the branch-and-bound method, improved by the application of a multi-modular graph convolutional network (MGCN) and deep Q-network (DQN), is employed to solve the maximum common subgraph (MCS) problem, resulting in the rapid and efficient matching of schematic and physical data.

Keywords: object detection; multi-modular graph convolutional network; deep Q-network; maximum common subgraph; branch-and-bound method



Citation: Cao, W.; Chen, Z.; Wu, C.; Li, T. A Method for Matching Information of Substation Secondary Screen Cabinet Terminal Block Based on Artificial Intelligence. *Appl. Sci.* **2024**, *14*, 1904. <https://doi.org/10.3390/app14051904>

Academic Editor: Sungho Kim

Received: 15 December 2023

Revised: 19 February 2024

Accepted: 21 February 2024

Published: 26 February 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

As a pivotal component of the electrical power system, substations' reliable and efficient operations are crucial to ensuring the electricity industry's safety and stability. Substations' primary equipment manages power generation, transmission, distribution, and consumption. Meanwhile, secondary equipment focuses on measuring, monitoring, controlling, and protecting the primary apparatus. In the smart grid era, substations have significantly advanced by integrating advanced sensing and artificial intelligence technologies beyond traditional infrastructure. These advancements facilitate functionalities such as automatic data collection, intelligent measurement, remote control, and real-time monitoring, further enabling the adaptation of control strategies within the station, based on the power grid's operational status. Aligning with the dual-carbon objectives, the national grid has initiated proposals to enhance the development of green and low-carbon technologies across all grid planning, construction, operation, and maintenance aspects, aiming to boost the energy sector's digital economy growth [1].

Constructing, inspecting, and maintaining secondary screen cabinets in substations require a thorough understanding of the diagrams and physical terminal blocks' elements and interconnections. Despite this, manufacturing, design, construction, and inspection data carriers have remained in unstructured forms like diagrams and documents for decades, leading to significant repetitive and inefficient manual data entry and duplication across various stages.

The rapid advancement of computer and communication technologies and artificial intelligence techniques, especially deep learning (DL), has reached new heights. These

technologies have attracted widespread attention and application in sectors such as transportation, electrical power, and communications [2–4]. As a result, digitally modeling secondary systems in substations, particularly protective screen cabinets, and rapidly comparing diagrams with their physical counterparts offer broad development prospects in the current construction, operation, and maintenance of substations.

Digital modeling's essence for secondary system screen cabinet diagrams and physical entities lies in object detection technology. In the last decade, DL-based object detection technologies have stood out, leading to the development of numerous remarkable models, such as the region-based convolutional neural network (R-CNN) [5], Fast-RCNN [6], and Faster-RCNN [7]. Notably, the You Only Look Once (YOLO) approach [8,9], treating object detection as a regression problem, has significantly advanced the field by segmenting images into grid units for regression and classification, thus outputting the categories and locations of entities [10]. Following extensive research, the model has evolved into YOLO version 7 (YOLOv7) [11], integrating efficient aggregation networks, auxiliary head detection, and model scaling mechanisms, thereby enhancing the YOLO series algorithms' speed and accuracy.

However, the level of object recognition for engineering diagrams and complex construction sites remains less than ideal. This challenge is due to the small size of micro-object targets and the complexity of effectively extracting their feature information, especially in terms of processing speed, noise resistance, and recognizing various characters and curves.

Moreover, intelligently operating and maintaining substations requires comparing digitally modeled diagrams and physical entities. This comparison, focusing primarily on node information, becomes increasingly complex as the scale of nodes expands, requiring significant computational resources. Identifying similar parts between two topological graphs is known as the maximum common subgraph (MCS) partitioning problem in graph theory [12], a challenge that has applications in software analysis, graph database systems, cloud platforms, and chemical drug synthesis [13–16].

Solving the MCS problem, an NP-Hard challenge, has always been difficult. The most effective current method is based on heuristic algorithms using branch-and-bound techniques [17], albeit time-consuming. Neural networks' advent has opened new pathways for addressing this problem. For instance, Liu et al. [18] integrated deep neural networks into the MCS framework, creating an MCS solution framework with a reinforcement learning (RL) network, which enhances the model's solving efficiency. Wang et al. [19] and Bai et al. [20] have also made significant contributions, applying deep learning networks to the MCS problem and making initial attempts to output MCSs directly using DL-based methods, respectively. Although the graph neural network (GNN) development has provided new insights, a universal model for solving this problem is still in development.

To address these problems, we propose a comprehensive framework for digital modelling and information matching of substation secondary panel cabinet terminal block diagrams and physical entities. The main contributions are as follows:

- To achieve intelligent recognition of secondary screen cabinet terminal block information, we design a multi-layer region segmentation algorithm based on You Only Look At Coefficients (YOLACT) and YOLOv7 with a multi-head self-attention mechanism and the space-to-depth module (YOLOv7-MS) to achieve the extraction of microelementary target regions in drawings. For the recognition of physical information of terminal blocks in the actual field, we also design YOLOv7 based on a differentiable binarization attention head (YOLOv7-DBAH) to solve the interference problem caused by shooting.
- To facilitate the digital information storage, we establish the network topology based on the Neo4j database for the diagram and physical system, respectively.
- To achieve fast matching of network topology data between terminal block drawings and physical entities, we design an improved branch definition method based on a multi-modular GNN (MGCN) and deep Q-network (DQN) to solve the MCS problem.

2. Methods

Before delving into the detailed methodologies employed in this study, it is essential to clarify the abbreviations and specific terminology used throughout this section to ensure a clear and comprehensive understanding for the reader. This study utilizes various advanced technologies and algorithms, including the multi-layer graph convolutional network (MGCN), deep Q-network (DQN), and maximum common subgraph (MCS), which are pivotal in achieving the objectives of intelligent matching and information extraction. The following subsections will detail the implementation and application of these technologies in the context of substation secondary terminal block information recognition, storage architecture, and topology comparison.

2.1. Overview

The overall framework for implementing intelligent matching of diagrams and physical information for substation secondary terminal blocks is shown in Figure 1. Initially, multi-layer object detection networks are employed to extract information precisely from both blueprints and physical entities. Subsequently, network topologies for both the diagrams and physical systems are established using the Neo4j database. This enables the digital storage of information related to the substation's secondary terminal systems. Finally, the MGCN and DQN are used to improve the branch-and-bound method. This approach is applied to solve for the MCS between the two topologies, with rapid and efficient matching of the diagrams and physical data being achieved as a result.

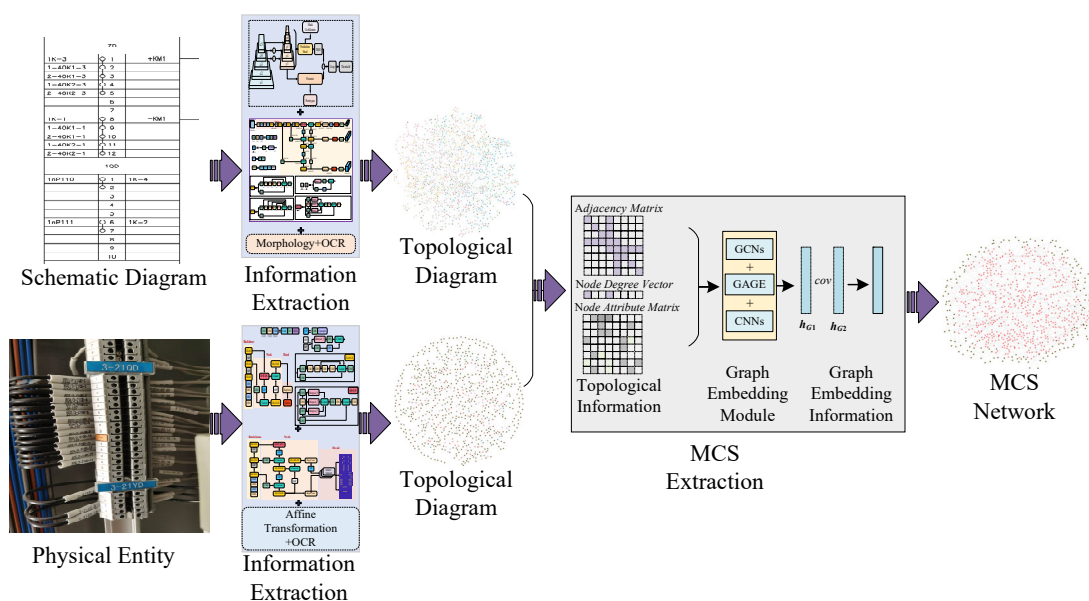


Figure 1. Overall framework for intelligent matching of substation secondary screen cabinet terminal block graphical and physical information.

2.2. Intelligent Recognition Technology for Secondary Cabinet Terminal Block Information

The construction, inspection, and maintenance of substation secondary cabinets rely on the complete understanding of the elements and their interconnections in the diagrams and physical entities. Traditionally, the connection line diagrams for substation terminal blocks are mostly in paper format, and the information about the physical terminal blocks is hard to quantify, which brings a huge workload to substation secondary system construction calibration and maintenance. Image recognition techniques are now increasingly being applied to actual production processes. The core task for the application of complex substation secondary system cabinet terminal blocks is to identify the connection relationships corresponding to each terminal block. The recognition of elements and connection relationships of each terminal block through machine vision will greatly enhance work

efficiency on-site and reduce labor costs. This chapter introduces the recognition technology for terminal block recognition technology for diagrams and physical entities, respectively.

2.2.1. Information Recognition Technique for Substation Terminal Block Diagrams Based on Regional Segmentation

The information related to connection relationships in equipment diagrams can be mainly categorized into two types: graphical element information and textual information. The electrical significance of graphical element information is the key foundation for analyzing the connection relationships of electrical diagrams, and the correct identification of the graphical elements is the key content of the study.

Considering the vast amount of information in terminal block diagrams, using detection models directly yields subpar results. This paper proposes a multi-layer region segmentation model to enhance the sensitivity of the detection network. The specific network structure is shown in Figure 2. The network model is divided into four layers, specifically as follows:

1. The terminal block area of an entire diagram is partitioned into single connected regions by the YOLACT network [21].
2. The small object detection module in the YOLOv7-MS is employed within the terminal block area for rapid detection of graphical elements and text positions in the diagrams.
3. Table characteristics within each single connected region are obtained through table detection methods in OpenCV [22], using morphological methods.
4. Textual information within different tables is recognized by utilizing the PaddleOCR model [23] in conjunction with a standard database.

In order to improve the performance of the detection head at the output side, focusing on the target objects expected to be detected, the multi-head self-attention mechanism (MHSA) in the BoT module [24] of the bottleneck converter is used between the ordinary convolutional layers at the output side. The MHSA achieves linear dimension uplift by full connectivity and implements self-attention in parallel for the four heads, and each self-attention learns multiple weights individually, with the feature weights being matrixed separately. Matrix multiplication with the feature weights is performed separately to obtain the corresponding feature output information. Finally, the result information of each head is spliced together to learn the feature fusion information from multiple perspectives. At the same time, it also has the feature of global attention on the 2D feature map, which organically correlates the features of the two phases of network extraction and fusion, and integrates the information under global attention to achieve better detection results.

In order to improve the detection effect of the microelement target elements of drawings, snowflake point deconvolution (SPD) [25], which is known for detecting low-resolution and small objects, is introduced at the output end, and the original image is resized using the frame-loop superscalar transformation technique. The SPD convolutional block consists of a spatial-to-depth layer (SPDlayer) and a non-strided layer. The SPDlayer is responsible for converting the original feature maps into intermediate feature maps with feature discriminative information.

The SPD convolutional building block firstly splits the original image of the drawing microdot target element into sub-feature maps, then splices the sub-feature maps into the intermediate feature maps, extracts the feature discriminative information therein, and finally filters and learns the extracted feature discriminative information through the filter. The above work makes the detection head more accurate in identifying the micrometric targets, and the introduction of SPD effectively improves the algorithm's performance in detecting the micrometric target elements of drawings.

In the improved network, the MHSA is added at the output to enhance the performance of the detection head, and the SPD convolutional building block is introduced to improve the detection of small, low-resolution objects.

Based on the aforementioned recognition results and the actual significance of the connection relationships within the tables, a comprehensive extraction of the schematic information can be performed.

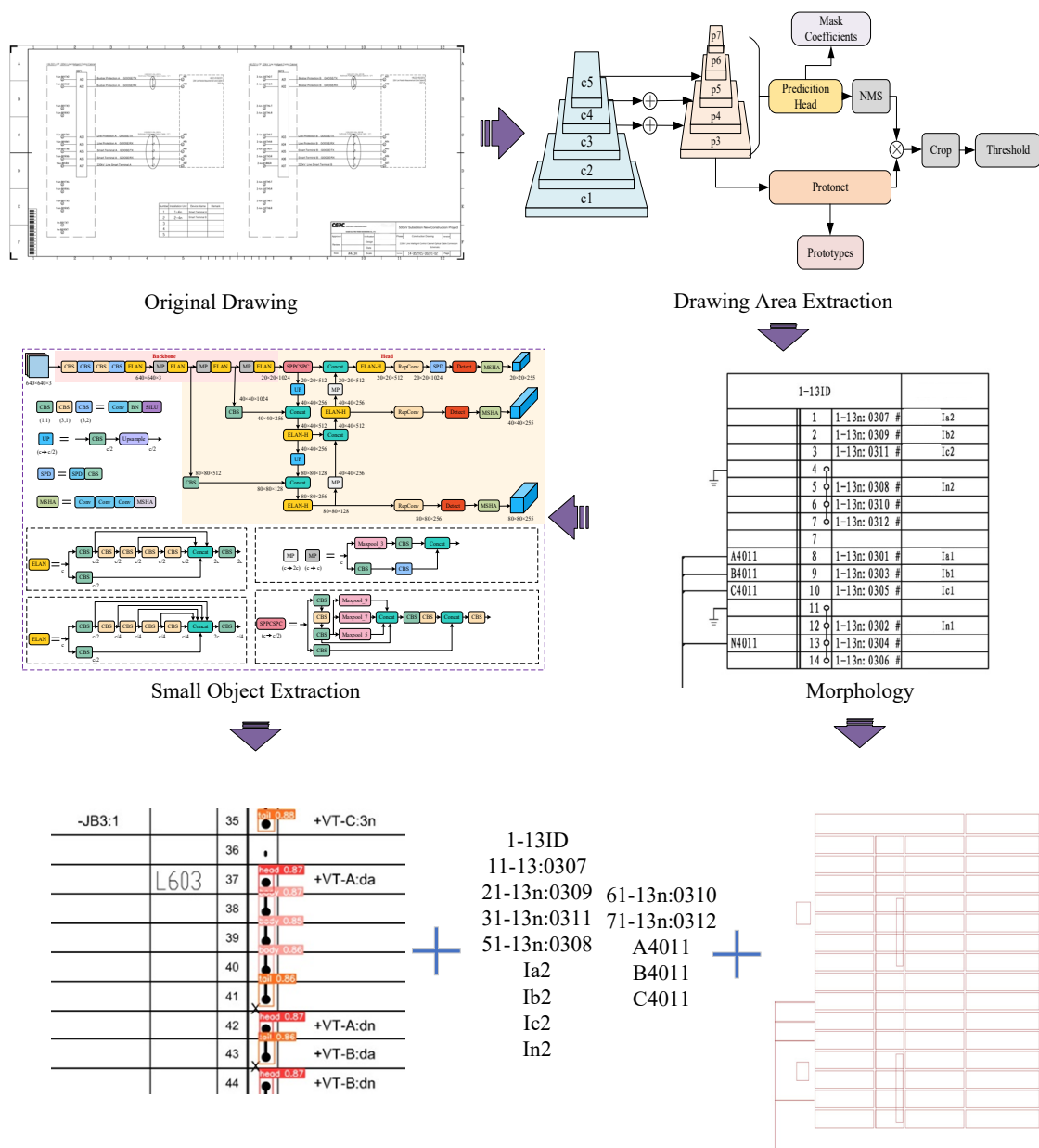


Figure 2. Framework for diagram identification of substation terminal block.

2.2.2. Recognition Technique for Substation Terminal Block Physical Information

Due to the significant influence of shooting angles and the background of the equipment, the identification of the physical information of the terminal block is more difficult compared with the identification of the drawing information, which cannot be used directly in the aforementioned model, based on which, this section designs a four-layer model for the identification of the physical information of the terminal block [26]. The structure of each layer and the function of the response are as follows (Figure 3):

1. The YOLOv7 area-oriented (YOLOv7-AO) model is used to pre-extract the target area of the physical area of the terminal block. The compacted and improved YOLOv7-AO model quickly locates the area where the target components such as the terminal block are located, assists the subsequent component detection model to locate the small

- target components, excludes the negative interference of the irrelevant background, and improves the overall recognition efficiency of the physical identification.
- The YOLOv7-DBAH model, which is designed for the detection of small target physical elements in terminal blocks, is used to locate all the boundary points of physical targets under the deformation of the shooting viewpoint.
- According to the coordinates of the original boundary points of the physical components, based on the principle of affine transformation, the affine parameters of different components are estimated, and the correction of all components is completed to restore the original shape of the component text.
- The PaddleOCR model continues to be used as the core algorithm for component text recognition in the aforementioned electrical context. The recognition and extraction of nameplate information of physical components are completed, serving as the unique identifier for components of the same type. The integration of image content in physical recognition is also completed.

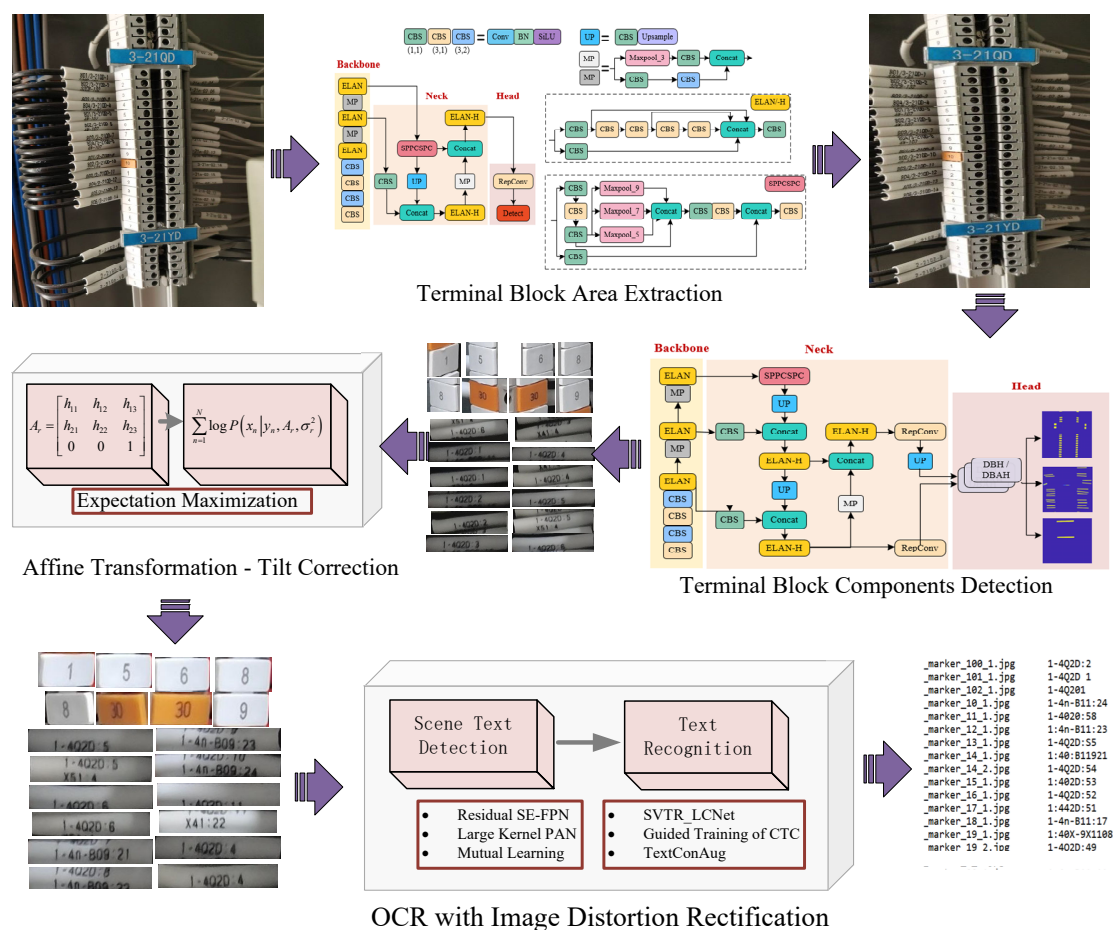


Figure 3. Framework for physical identification of substation terminal block.

The DB detection head transforms the bounding box regression subproblem in the standard target detection problem into a pixel-level classification problem, which can obtain a more close fit to the actual boundary contour of the physical object, and better solves the deformation problem in the target image studied in this chapter. In order to solve the multi-target class synchronous detection problem better, the DBAH detection head is constructed on the basis of the basic DB detection head by combining the spatial and channel attentional weighting parts to improve the model detection performance.

The DBH detection head part has a simple structure, where the two sets of feature maps are directly spliced and merged together to generate the probability map and threshold

map for target class component identification through a 3×3 convolutional layer (ReLU activation function) and a 1×1 convolutional layer (Sigmoid activation function), and the final target binarization map is processed by the preset differentiable binarization function to output the final target binarization map. Individual target instances of the same type can be identified by a basic boundary extraction algorithm that determines the exact polygonal boundary contours and approximates them using a rotating rectangular box to output a uniform format of target location information to be recognized.

The DBAH detector head adds the feature map spatial attention mechanism and the channel attention mechanism to the DBH detector head structure to strengthen the local key information of the corresponding features of different detector heads of each class. Before splicing and merging the two sets of feature maps, firstly, the location weights of the feature maps related to the target class to be predicted are superimposed to the original features by independent spatial attention mechanisms; secondly, the attention vectors of the two sets of feature maps are spliced together by the channel attention mechanism and then further computed using the convolution layer to compute the global channel attention vectors further; thirdly, the spatial attention-enhanced feature maps are spliced and fused with the replicated channel attention vectors by dimension; finally, the idea of DBH detection is continued, and the two convolutional layers are used to convert the feature maps to target probability maps, threshold maps, and binarized maps, and to obtain the target location information.

The terminal block table area is a crucial part of the terminal block diagrams, containing the main components of the terminal block equipment, such as terminal names, terminal numbers, number tubes, test terminals, and shorting links. The key to identifying effective information in terminal block diagrams accurately lies in recognizing the structure of the terminal block table area, including the coordinates of the cells within the table and their row–column relationships. Therefore, this paper employs a table detection algorithm based on grayscale binarization and morphological operations to identify the table structure.

Ground symbols are located outside the terminal block table area and occupy a very small pixel ratio in the terminal block's single connected area, falling into the category of micro-element target detection. Considering the characteristics of the graphics and regional features, a dynamic threshold sliding window segmentation based on aspect ratio is adopted, setting the overlap step length to be greater than the maximum pixel of a complete ground symbol to ensure that each segmented image contains at least one complete graphic element. The segmented new images are fed into the YOLOv5_groundconnect network for ground symbol recognition. The results are then restored to their coordinates in the unsegmented terminal block single connected area through a coordinate restoration algorithm. Subsequently, a deduplication algorithm is applied to correct the same ground symbols identified in different segmented images, eliminating the quantity error caused by repeated recognition.

Shorting links, located within the terminal block table area, come in five forms and have a wide range of lengths, making direct detection using deep learning object detection challenging. However, the head, middle part, and tail of the shorting links have fixed shapes and lengths, conforming to the precision rules of the deep learning algorithm for the features of the target to be learned. Based on this, a discrete-normalization algorithm is proposed, discretizing the shorting links into three parts: head, body, and tail. Since all shorting links are distributed within the cells of the terminal block table area, a dynamic threshold sliding window segmentation algorithm based on table row–column relationships is used, performing segmentation every ten rows per cell without setting an overlap step length, eliminating the need for deduplication of shorting links. The segmented images are fed into the YOLOv5_shortconnectedpiece network for the recognition of discrete shorting links. The recognition results are for the shorting links in a discrete state, which need to be restored to the unsegmented terminal block table area through coordinate restoration, followed by normalization of the shorting links to generate the final recognition results.

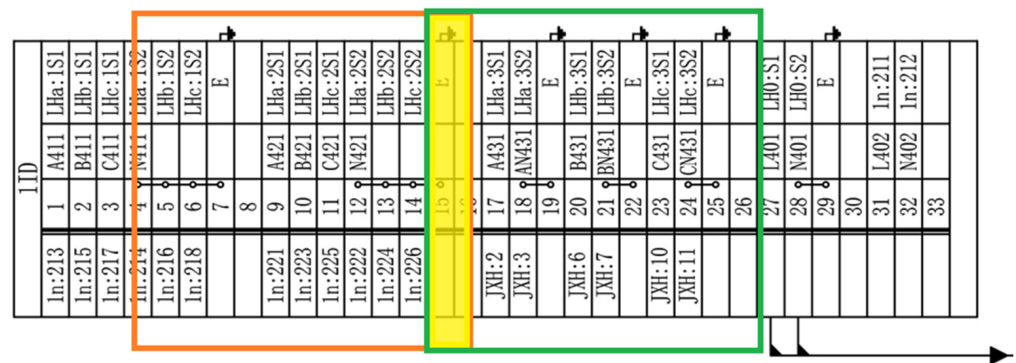
2.3. Architecture for Storing Secondary Cabinet Element Information Based on Neo4j Graph Database

This section outlines the innovative approach adopted for structuring and storing the intricate details of secondary cabinet terminal blocks within substations. Utilizing the Neo4j graph database, we establish a robust and dynamic architecture that not only accommodates the complex relationships between various terminal block components but also facilitates efficient data retrieval and analysis. The subsequent subsections detail the methodology for constructing a comprehensive topological model that represents both the schematic and physical aspects of terminal block configurations, thereby laying a foundational framework for intelligent data management and application in substations.

2.3.1. Construction of Topological Information for Secondary Cabinet Diagrams

To construct a model for secondary cabinet terminal block schematic information based on a graph database, it is necessary to map the relational model to a graphical network model with a $G(V, E)$ structure, where V represents the set of vertices and E represents the set of edges. Each graphical element and its corresponding connection relationships being rationally mapped will facilitate the construction of the graph database and subsequent querying of graphical element information.

As shown in Figure 4, the graphical elements in the secondary cabinet terminal block diagrams include terminal block numbers, terminal connection entities, grounding symbols, external connection lines, and wiring terminals, among others.



considered separate graphical elements and are set as nodes in the model construction. The node name is represented by the text on the connection line, and its corresponding attribute is the connection line itself.

5. Based on the connectivity relationships among the various elements obtained in Section 3.2, corresponding connectivity relationships are established in the graph. Attributes for connections between adjacent nodes are subsequently formed.
6. Each node corresponds to different attributes (such as circuit type, whether the node belongs to a power source or a signal, etc.), which can serve as one of the input factors when constructing the graph database.

Based on the aforementioned principles and methods, Figure 5 illustrates the schematic topological diagram of the terminal blocks. In the figure, terminals from different rows are represented by distinct colors, and the functionalities of test terminals and connected entities of output interfaces are indicated as node attributes. The model only displays two types of connecting lines: one type constitutes the overall continuity of the terminal row, and the other type represents connecting lines with existing relationships. In reality, there are various forms of connecting lines, which can be further distinguished based on different types of circuits or signal types.

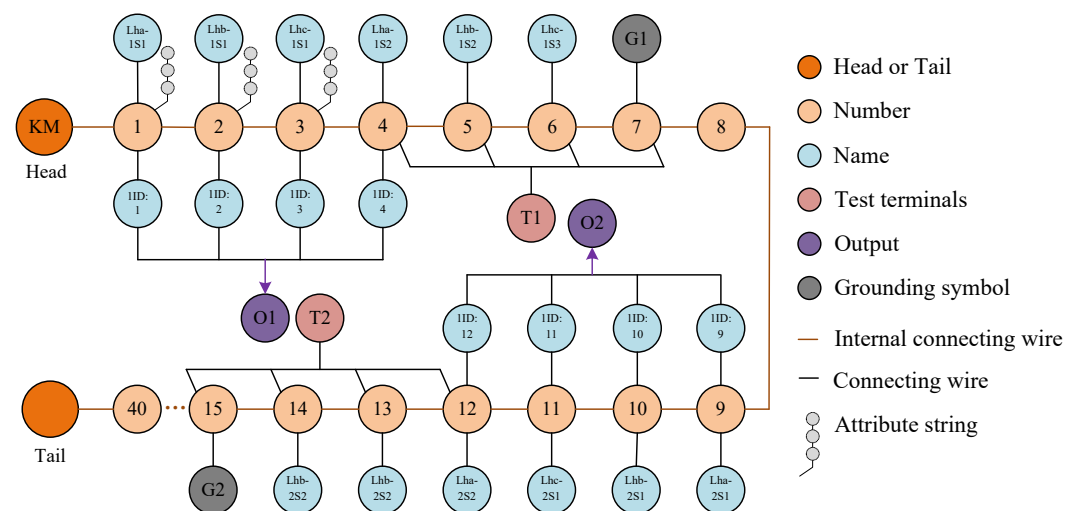


Figure 5. Connection model of the schematic diagram of secondary screen cabinet terminal block.

Comprehensive information construction for the secondary cabinet terminal block diagrams is carried out based on the aforementioned methods to form the topology map of the secondary cabinet diagrams. Due to the diversity and complexity of the terminal block equipment in the substation's secondary cabinets, there are significant differences between different equipment diagrams in terms of node scale and node connection types.

2.3.2. Construction of Topological Information for Secondary Cabinet Entities

As shown in Figure 6, the physical elements of the secondary cabinet terminal blocks only include the terminal block headers, terminal sequence numbers, and external terminal row names. Therefore, the constructed image-based model is relatively simple, with the following specifics:

1. Each terminal block header serves as the unique identifier for that block. The name of each terminal on the block is determined by its corresponding sequence number and block name, and it corresponds to a node when forming the topological diagram.
2. The connecting lines on both sides of each terminal block number are identified to form corresponding nodes.
3. The attributes corresponding to each node are stored in the model as attribute strings.

Based on the aforementioned principles and methods, Figure 7 illustrates the schematic topology map of the terminal block diagrams for the secondary cabinet's physical information. From the figure, it can be seen that the topological diagram formed by the physical terminal blocks is much simpler than the schematic topological diagram. This is due to the fact that the amount of information in the diagram of the compression rate is higher, and each node needs to be well known for all of its functions and the direction of the connection.

In reality, after merging the nodes of the schematic and physical topologies for the entire substation, the information in these two types of topological diagrams for the same substation should be consistent. Figure 12 displays the schematic and physical topological diagrams of a small-scale substation.



Figure 6. The physical entity of secondary screen cabinet terminal block.

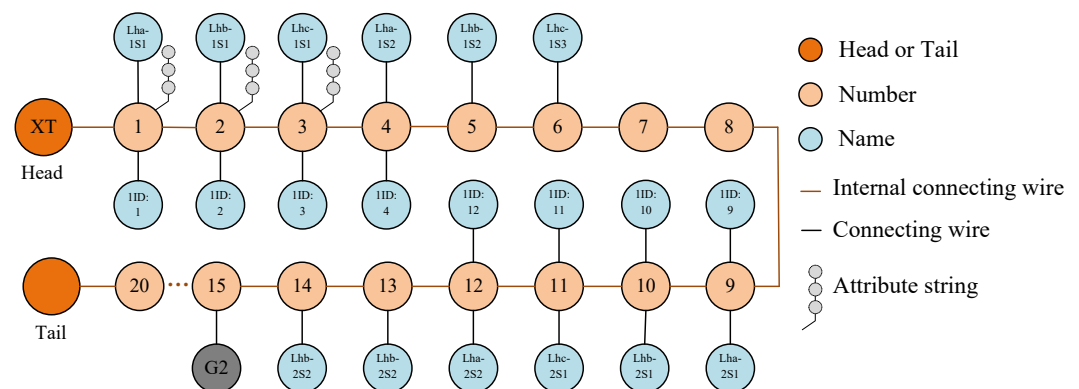


Figure 7. Connection model of the physical entity of secondary screen cabinet terminal block.

2.4. Improved Branch-and-Bound Algorithm for Topology Comparison between Diagrams and Physical Entities

Based on the methods in Section 2.3, large-scale schematic and physical topological diagrams of substation terminal blocks can be obtained. Some of the crucial tasks in substation construction and inspection are to assess the connectivity between the diagrams and physical entities accurately, and to identify any discrepancies that deviate from the design specifications. To achieve this, the use of reasonable ways to compare the two topological relationships can be obtained in the topological map of the nodes, to check for

nodes that do not match, thereby facilitating the evaluation of construction outcomes. This process is essentially a search for the largest common subgraph of the two, belonging to the typical topological problems.

2.4.1. MCS Algorithm Based on Branch-and-Bound Method

Since the MCS is identified as an NP-hard problem, the most effective current method for its identification between two topological graphs is based on a heuristic algorithm known as the branch-and-bound method.

The branch delimitation algorithm is essentially a depth-first algorithm. For two topological graphs G_1 and G_2 , the algorithm starts with an empty subgraph and iteratively adds a pair of nodes in a specific manner, ensuring that the added pair still maintains the topological consistency of the subgraph. During each search iteration, the current search state is represented by a_t , consisting of elements from both topological graphs. The pair (i, j) represents the node numbers corresponding to the two subgraphs in each search step, where node i is the newly added node in G_1 and node j is the newly added node in G_2 , denoted by b_t . Decide whether the current state is to continue the search or return to the parent search node based on the results of the impact of the added node. Different search strategies can be employed for the search of new nodes, and these strategies influence both the efficiency and the focus of the algorithm. For instance, in [17], the degree of the node is used as a criterion for determining the priority of the search node. The higher the degree of the node, the more likely it is to be the next node.

In fact, at each step of the search, it is possible to evaluate whether the nodes searched at each step satisfy the requirements by calculating the bounding value $Bound_t$, and, for this purpose, the concept of “domain” is introduced, which divides the remaining nodes into several equal parts. In the set R , which consists of all domains in a given state, $R_k = [V_1, V_2]$ is composed of two sets of nodes. In this context, V_1 and V_2 have the same connection patterns for nodes that have already been matched. To ensure the single-connectivity of the subgraph, the range of node pair selection, or the action space, A_t , is confined to the domain, $R^{(c)}$, that is connected to the already matched nodes. Based on the definition of domain and according to [12], the calculation method for $Bound_t$ is $\sum_{R_k \in R} \min(|V'_{k1}, V'_{k2}|)$.

2.4.2. Node Search Model Based on the DQN with MGCN Networks

As previously mentioned, different node search strategies yield varying results in the MCS search algorithm. Traditional algorithms often suffer from computational overload and are ill-suited for handling large-scale topological structures in real-world scenarios. For this reason, this paper adopts the form of fusion multi-module graph convolution for the extraction of topological graph information, and then realizes the evaluation of the results of different search nodes through the deep Q-network, so as to achieve the fast implementation of the subgraph search algorithm.

Considering that the action space of each new search node may be large and has many influencing factors, it is difficult to achieve a quantitative description of each action result. Instead, using the DQN, each node action space (a_t, b_t) can be mapped to a specific score, $Q(a_t, b_t)$, through the automatic learning of various neural network parameters, which in turn achieves the autonomous selection of node pairs.

Based on the aforementioned approach, in order to consider the matched nodes and the whole graph information fully, the graph convolution network is used to embed the node subgraph information in the topological graph that has been matched, which are denoted as h_{a1} and h_{a2} , respectively. The information of the whole topological graph is denoted as h_{G1} and h_{G2} through the embedding network. Each state can thus be represented as $h_{st} = \text{CONCAT}(h_{a1}, h_{a2}, h_{G1}, h_{G2})$. The node information corresponding to the action space can be represented as $h_{bt} = \text{CONCAT}(h_i, h_j)$ after passing through the embedding

function. The corresponding Q-value (Equation (1)) can be represented by the MLP network as follows:

$$\begin{aligned} Q(a_t, b_t) &= \text{MPL}(\text{CONCAT}(h_{at}, h_{bt})) \\ &= \text{MPL}(\text{CONCAT}(h_{G1}, h_{G2}, h_{a1}, h_{a2}, h_i, h_j)) \end{aligned} \quad (1)$$

To enhance the impact of adding node pairs (i, j) on the search result further, a discount factor, λ , is introduced. The Q-function can be further improved, $Q^*(a_t, b_t) = 1 + \lambda V^*(a_{t+1})$, where V represents the reward function. To introduce symmetry into the model, a function $IET(\cdot)$ with commutative properties is introduced, enhancing the model's adaptability to symmetric graphs. Using the previously defined domain to amplify the impact of different actions on the results further. For domain D^k , its embedded information can be represented as Equation (2), where $GETOUT(\cdot)$ denotes the aggregation of all elements in a particular set to obtain the result.

$$\begin{aligned} h_{R_k} &= IET(GETOUT(\{h_i \mid i \in V'_{k1}\}), \\ &\quad GETOUT(\{h_j \mid i \in V'_{k2}\})) \end{aligned} \quad (2)$$

The embedding information for all connected domains, h_{R_c} , can be expressed as $h^{Rc} = GETOUT(\{h_{R_k} \mid k \in R^{(c)}\})$.

Based on the above definition, the Q function is rewritten as shown in Equation (3).

$$\begin{aligned} Q(a_t, b_t) &= 1 + \lambda \text{MPL}(IET(h_{G1}, h_{G2}), \\ &\quad IET(h_{a1}, h_{a2}), h_{Rc}, h_{R0}) \end{aligned} \quad (3)$$

The DQN network employed in this study is illustrated in Figure 8. For graph embedding of topological nodes, we select adjacency matrix A , node attribute matrix B , and node degree vector l as inputs to the graph embedding module. The GCN is used for feature extraction from both the entire topological graph and its matched subgraphs. To enhance the efficiency of graph embedding, we adopt the SAGE model described in [27] for rapid embedding of node and graph features. This is followed by a CNN to form embedding information vectors for different topological graphs, and convolutional computations are performed to enable the function $IET(\cdot)$ for topological graph information. For node embedding, we opt for a graph embedding model comprising a GCN, attention mechanism, and CNN, as proposed in [28]. These three feature extraction networks facilitate the extraction of previously mentioned embedding information vectors h_G , h_s , h_{D_c} , and h_{D_0} . The CONCAT function is used to concatenate these vectors into a global feature vector, which is then further abstracted to a one-dimensional form through a MLP. Finally, activation functions are applied to map different action spaces to DQN scores.

In the actual training process, the overall loss function is represented by Equation (4). Here, y_t represents the target for iteration t , and $Q(a_t, b_t)$ is the predicted score. In this case, y_t represents the remaining size of the largest common subgraph from a_t to its leaf nodes in the current branch of the search tree.

$$\text{loss} = (y_t - Q(a_t, b_t))^2 \quad (4)$$

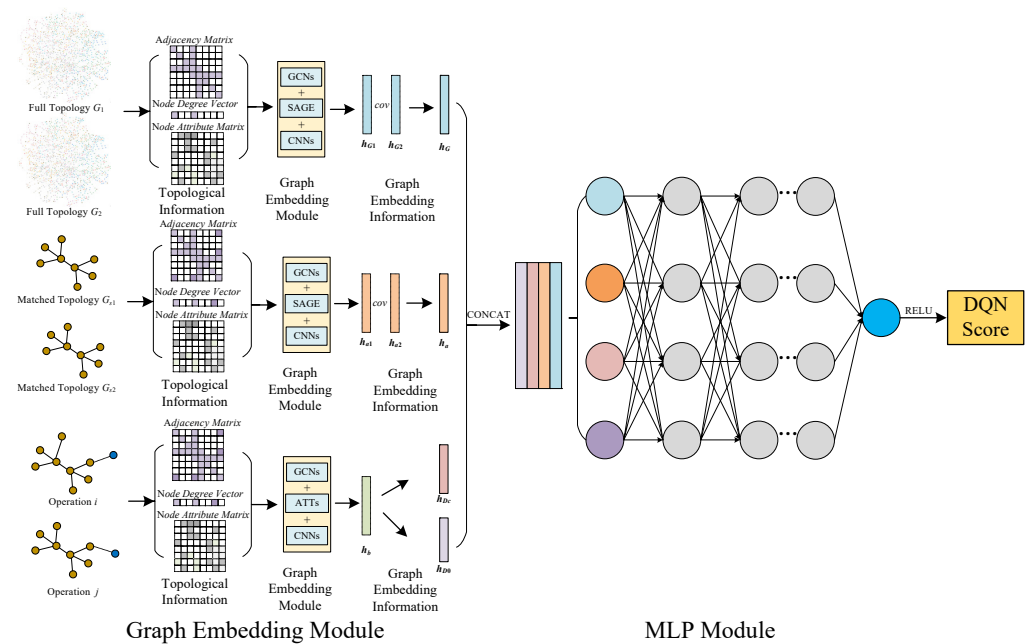


Figure 8. MGCN and DQN network structures.

3. Analysis of Experimental Results

This section delves into the comprehensive evaluation of the experimental outcomes derived from the implementation of our proposed methodologies for intelligent matching and recognition of secondary cabinet terminal block information. Through a series of meticulously designed experiments, we assess the performance of our models across various dimensions, including their ability to recognize diagrammatic representations and physical entities of terminal blocks accurately, as well as their effectiveness in matching these two forms of data. The following subsections provide a detailed account of the computational resources utilized, the dataset characteristics, the evaluation metrics adopted for performance assessment, and a thorough analysis of the results obtained, thereby illustrating the practical applicability and efficiency of our approach in real-world substation environments.

3.1. Computation Platform

Table 1 shows the hardware and software of the computation platform.

Table 1. Computation platform.

Hardware/Software	Version
Operating System	Windows 10
Memory	32 GB
CPU	AMD Ryzen 9 5900X 12-Core
GPU	NVIDIA GeForce RTX 3080
Python	3.8
Pytorch	1.9
PaddleOCR	2.6
Neo4j Desktop	1.5.8

3.2. Dataset

Currently, there is a lack of a publicly available image dataset dedicated to the identification of terminal block components. Therefore, it is essential for future detection models to create a dataset that is both of high quality and large in size. This dataset is carefully compiled using high-resolution cameras, and manual annotation is conducted using the

Labelling tool. The statistical findings pertaining to the dataset are exhaustively detailed in Table 2.

Table 2. Dataset.

	Train	Valid	Test
Schematic Diagram	825	137	137
Physical Entity	1306	194	194

3.3. Experiment Settings

To extract the terminal block area, we utilize stochastic gradient descent (SGD) with a momentum factor set at 0.9 for optimizing the network. The initial learning rate is established at 0.01, accompanied by a weight decay parameter of 0.0005. We configure the training process to span 270 epochs, processing data in batches of 16.

3.4. Evaluation Metrics

In this study, a range of metrics is utilized to evaluate the performance of various models in detection tasks thoroughly. Precision (5), recall (6), and F1 score (7), derived from four fundamental statistical parameters—true positive (TP), true negative (TN), false positive (FP), and false negative (FN)—are the primary metrics for this assessment. Precision is an indicator of the accuracy of positively identified samples among all detected samples, whereas recall indicates the proportion of accurately identified positive samples out of all actual positives. The F1 score, representing the weighted balance between precision and recall, is the harmonic mean of these two metrics. Additionally, mean average precision (mAP) (8), a prevalent metric in object detection, is also considered. mAP evaluates the model's average efficacy across all categories at varying levels of confidence.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (7)$$

$$\text{mAP} = \frac{1}{Q} \sum_{q=1}^{\ell} \int_0^1 P_q(R) dR \quad (8)$$

3.5. Analysis of Information Recognition Effectiveness for Secondary Cabinet Terminal Blocks

In this section, we embark on a detailed examination of the effectiveness of our proposed model in recognizing and matching information related to secondary cabinet terminal blocks within substations. By leveraging advanced object detection algorithms and sophisticated matching techniques, we aim to demonstrate the model's capability in accurately identifying both diagrammatic and physical representations of terminal blocks. This analysis encompasses a comprehensive evaluation of the model's performance across various scenarios, including the recognition of diagrams and physical entities, as well as the matching of these elements to ensure consistency and accuracy in the representation of substation secondary systems. The insights gained from this analysis are crucial for validating the practical utility and efficiency of our approach in enhancing the operation and maintenance of substations.

3.5.1. Analysis of Diagram Recognition Effectiveness for Secondary Cabinet Terminal Strips

The model proposed in Section 2.2 is employed for diagram recognition. The fully trained model is applied to new schematic recognition tasks, and its performance is evaluated based on the recognition of graphical elements within the diagrams. Figure 9 illustrates the model's recognition results for graphical elements such as grounding symbols, terminal connectors, and cells.

The results of the ablation experiments are shown in Table 3. The YOLOv7+Mash+SPD used in the algorithm of the article improves the mAP value by 25.7%, 10.6%, and 8.9%, respectively, compared with the initial YOLOv7, YOLOv7+Mash, and YOLOv7+SPD, and, comprehensively, the algorithm of this article adopts the combination of the integration of Mash's self-attention mechanism and the SPD convolutional building block. The improvement effect is the best, and its detection effect is better than other combinations in the complex scenario of end-row drawing microelement targets, especially compared with the baseline YOLOv7, to detect the drawings, as the mAP value is significantly improved by 25.7% points, and the leakage rate is reduced by 61.5% points.

Table 3. Ablation experiment with different improved modules.

Model	Precision	mAP	Recall	False Negative Rate
YOLOv7	0.795	0.653	0.712	0.288
YOLOv7+Mash	0.852	0.742	0.875	0.125
YOLOv7+SPD	0.861	0.754	0.878	0.122
YOLOv7+Mash+SPD	0.887	0.821	0.889	0.111

As can be seen from Figure 9, the model proposed in this study is capable of accurately recognizing the graphical elements in the diagrams. To validate the effectiveness of the algorithm further, metrics such as mean average precision (mAP), recall, and false negative rate are employed as evaluation criteria for the model's performance. The model's performance is compared with traditional object detection algorithms such as SSD [29], R-CNN, YOLOv5 [10], and conventional YOLOv7 to demonstrate its effectiveness. The specific results are presented in Table 4.

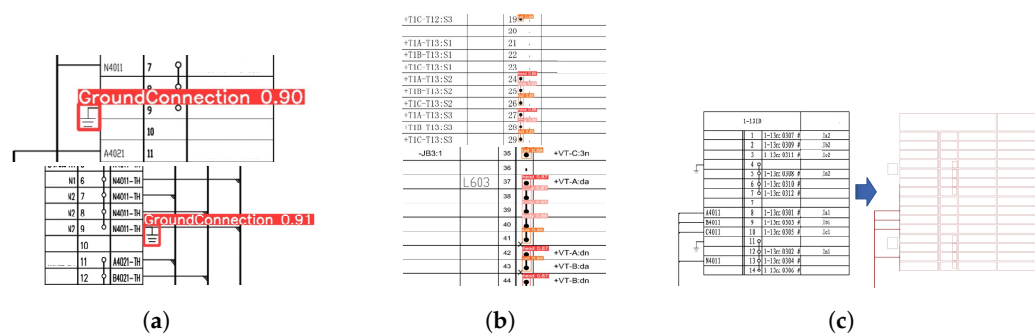


Figure 9. Performance of the proposed model on different entities in schematic diagrams. (a) Grounding Symbols. (b) Grounding Symbols. (c) Cells.

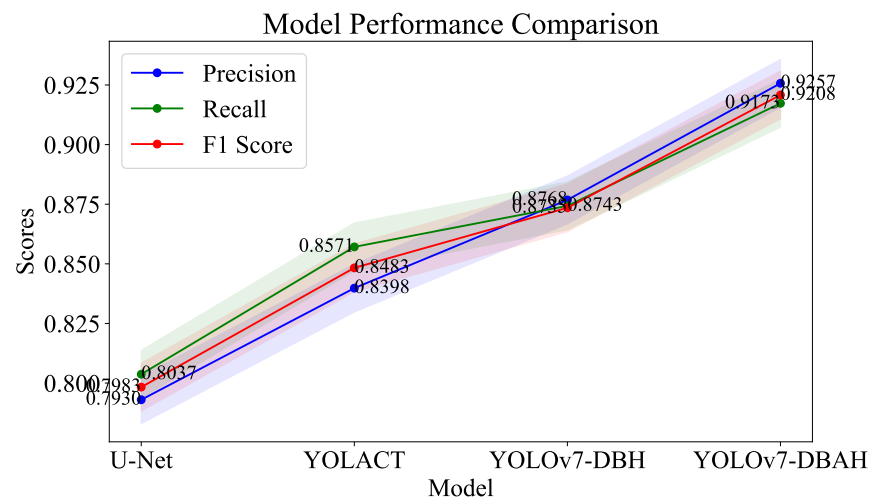
As can be seen from Table 4, when compared with traditional algorithms like SSD, R-CNN, YOLOv5, and conventional YOLOv7, the proposed algorithm in this study shows a significant improvement in mean average precision (mAP) of 70.3%, 48.2%, 30.5%, and 25.7%, respectively, and a reduction in the false negative rate of 79%, 75.4%, 64.8%, and 61.5%. This indicates that the improved algorithm effectively enhances the detection capability for small entities.

Table 4. Performance of the proposed model on different entities in schematic diagrams.

Model	Precision	mAP	Recall	False Negative Rate
SSD	0.492	0.482	0.473	0.527
R-CNN	0.598	0.554	0.549	0.451
YOLOv5	0.719	0.629	0.685	0.315
YOLOv7	0.791	0.653	0.712	0.288
Our	0.887	0.821	0.889	0.111

3.5.2. Analysis of Physical Entities Recognition Results for Secondary Screen Terminal Blocks

In Section 2.2, we introduced a model specifically designed for recognizing the physical entities of terminal blocks. To evaluate the target detection performance of the YOLOv7-DBAH model, we conduct comparative experiments using the same test data. The results, which encompass a comprehensive comparison of category indices, are presented in Figure 10 and Table 5.

**Figure 10.** Comparison of average results of different models.**Table 5.** Comparison of average results of different models.

Model	Precision	Recall	F1 Score
U-Net	0.7930	0.8037	0.7983
YOLACT	0.8398	0.8571	0.8483
YOLOv7-DBH	0.8768	0.8735	0.8743
Our	0.9257	0.9173	0.9208

The experimental outcomes demonstrate that the YOLOv7-DBAH model excels in average index results across three types of targets, with its lowest indicator surpassing 0.91. This performance places it distinctly ahead of the YOLOv7-DBH and YOLACT models, which form the second tier, with index values around 0.85. The conventional U-Net segmentation network lags slightly behind, maintaining an evaluation index near 0.80. These findings indicate that the YOLOv7-DBH structure, as developed in this study, outperforms baseline segmentation networks in detecting and identifying terminal strip-related physical components. Moreover, the YOLOv7-DBAH model, which integrates spatial and channel attention mechanisms, enhances the intermediate feature map with a higher proportion of target recognition-related information. This enhancement enables different detection heads to learn features pertinent to the target, culminating in superior physical component recognition.

Analyzing the sub-index results for the three component types, including terminal numbers, reveals that the ratio of positive to negative pixel areas in the sample significantly influences the precision and recall calculations. Even with the same recognition model, the terminal number and number tube, benefiting from a larger target count, demonstrate a marked improvement in accuracy compared with the terminal inscription target. However, the recall index results show that the DB-type detection head brings the terminal inscription target's performance closer to that of the other two target types. This suggests that the DBAH detection head, designed with an attention mechanism, tends to enhance the feature extraction of target types in imbalanced samples. Despite this, the YOLOv7-DBH model does not outperform in recognizing terminal inscription targets in the F1 score comparison. However, its mean results slightly exceed those of the YOLACT model, primarily due to better performance with terminal number and number tube targets. The YOLOv7-DBAH model, augmented with the attention mechanism, shows improvements across all indicators for the three types of physical component recognition. It not only raises the precision and recall rates, which were already higher than the YOLACT model, to above 0.91 but also enhances the F1 score for terminal inscription targets from 0.83 to over 0.89. This significant improvement reaffirms the efficacy of the DBAH detection head in enhancing target recognition.

An example of the recognition results is shown in Figure 11, based on an actual photograph of a terminal block. The original photo in the figure is significantly affected by the background. After area localization using YOLOv7-AO, irrelevant background areas and severe distortions and obstructions on the left side of another terminal block are effectively eliminated. In the second layer of the model, DB is introduced. The YOLOv7-DBAH model, tailored for detecting small physical entities of terminal blocks, is used to locate all physical object boundary points under distorted shooting angles accurately, achieving good results. Finally, perspective technology is used to correct the distorted text information, enabling complete physical entities recognition.

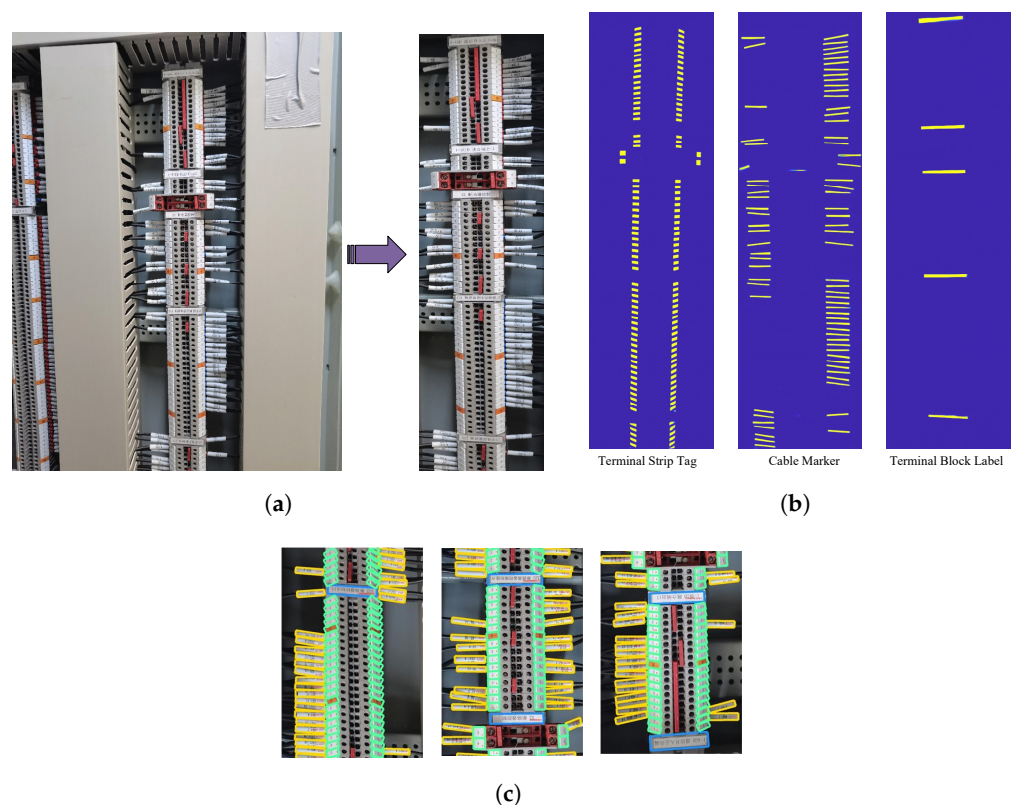


Figure 11. Results of the physical recognition model. (a) Area Location. (b) Outcome Binarization Maps. (c) Identification Results.

3.5.3. Analysis of Diagram-to-Physical Matching Results in Substation Secondary Systems

As previously mentioned, after the recognition model extracts data and merges nodes from both the diagrams and the physical entities of terminal blocks, two separate topological maps are formed for a substation (see Figure 12). By comparing the details of these two maps, the verification of connection lines in the substation's secondary system can be achieved. This section experimentally validates the effectiveness of the topological mapping algorithm proposed in Section 2.4 and compares it with traditional algorithms to demonstrate its superiority.

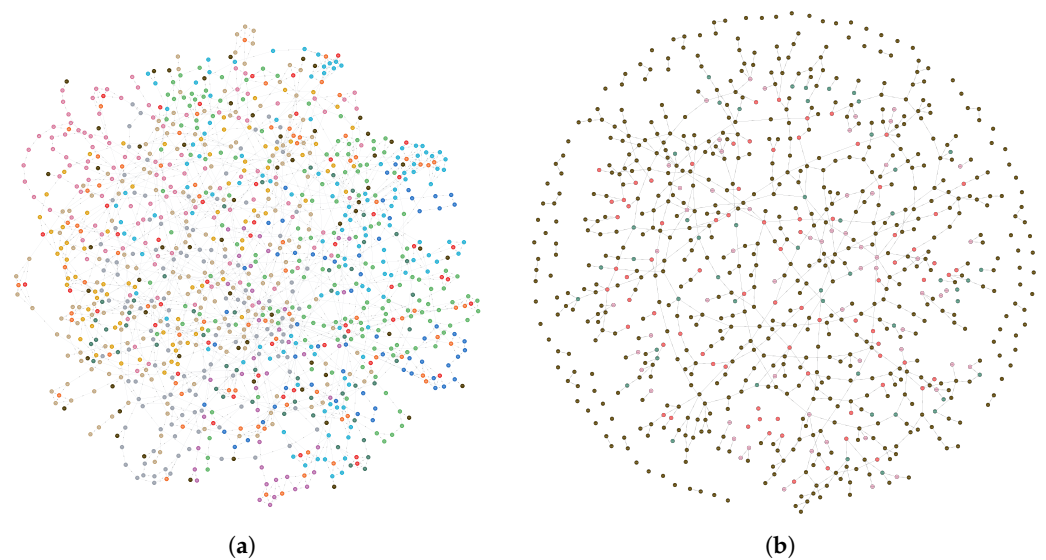


Figure 12. Visualization of topological information in substation secondary screen cabinet terminal block graphical and physical entities. (a) Topological Information of the Schematic Diagram. (b) Topological Information of the Physical Entity.

The node embedding and graph embedding networks within the algorithmic model are already detailed in Section 2.4.2. For the DQN network, a four-layer MLP network is chosen to map the graph-embedded information to corresponding scalars, which are then transformed into Q-scores through the RELU activation function.

Normally, the training of the model requires a substantial amount of topological map data, and consumes considerable time and computational resources. To mitigate this, pre-training is employed to complete the initial parameter optimization, significantly reducing the training cost of the model. For this purpose, the model is trained on a specialized network based on the graph embedding model trained in [4], significantly reducing the model's convergence time.

To validate the effectiveness of the model, this study compares its algorithm with traditional algorithms based on two metrics: the proportion of the largest common subgraph (R) and the algorithm execution time (t/s). The control groups are as follows: (1) the most basic point-to-point data comparison, denoted as N-to-N, which theoretically can fully realize node-to-node relationship comparison but is the most time-consuming; (2) precise MCS algorithms, such as MCSP+RL [18]; and (3) learning-based graph matching models, like GW-QAP [30]. Additionally, to validate the effectiveness of the DQN network designed in this study, a comparative experiment is added where a random number replaces the output value of the DQN network, denoted as DQN-rand. For the aforementioned five groups of models, algorithms are evaluated sequentially using networks of varying scales. The results are shown in Tables 6 and 7.

Table 6. Results of the MCS occupancy rate.

Size	N-to-N	GW-QAP	MCSP	DQN-Rand	Our
Node: 1000					
Edge: 2130	0.528	0.419	0.411	0.462	0.469
Node: 10,000					
Edge: 21,700	0.501	0.402	0.323	0.302	0.462
Node: 50,000					
Edge: 43,200	0.515	0.395	0.301	0.319	0.530
Node: 100,000					
Edge: 586,200	0.533	0.389	0.222	0.213	0.533

Table 7. Results of execution time.

Size	N-to-N	GW-QAP	MCSP	DQN-Rand	Our
Node: 1000					
Edge: 2130	1022	539	252	201	189
Node: 10,000					
Edge: 21,700	5218	2213	938	235	247
Node: 50,000					
Edge: 43,200	7000+	4825	1932	934	237
Node: 100,000					
Edge: 586,200	7000+	6712	3694	4592	2623

As seen in Figure 13a, when the scale of the nodes is small, all algorithms perform similarly in terms of the R metric. The operational outcomes are all closely aligned with the N-to-N strategy. As the node scale increases, the performance of the DQN-rand, MCSP+RL, and GW-QAP models gradually deteriorates. These three models exhibit weak adaptability to large-scale topological graphs. When the Q-scores of the DQN network are replaced by random numbers, the optimization results cannot be fed back into the network input, causing the model to degrade to the original MCS framework. This further substantiates the effectiveness of the algorithm proposed in this paper. Note: the N-to-N strategy has not been fully executed when the node scale exceeds 50,000.

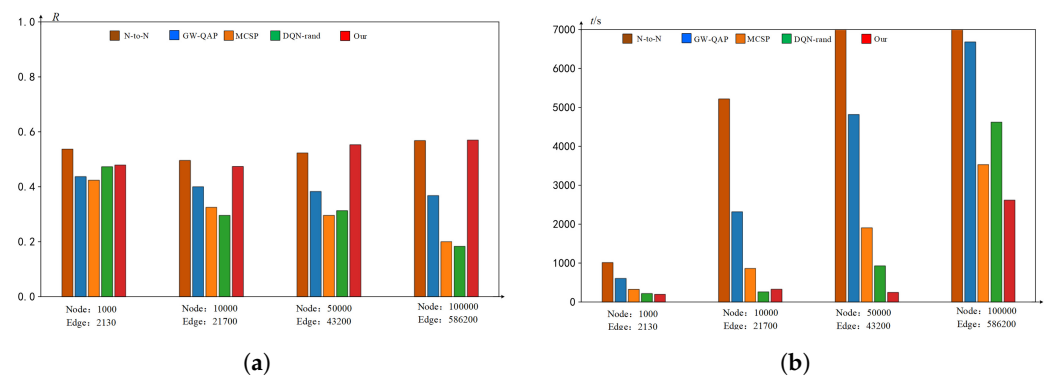
**Figure 13.** Comparison of different algorithm execution results. (a) Results of the MCS Occupancy Rate. (b) Results of Execution Time.

Figure 13b illustrates the time taken to complete different tasks. It is evident that, for the N-to-N strategy, the execution time increases rapidly as the complexity of the network structure grows. When the network scale reaches 50,000, it essentially becomes infeasible to complete the data comparison tasks. For large-scale topological structures, the model proposed in this paper has the shortest running time. This is because the node information embedding model chosen in this paper is the SAGE model, which is well-suited for embedding information in large-scale topological graphs.

Figure 14 provides a process diagram of the model executing the comparison between the schematic and the physical layout of the substation's secondary system terminal strip.

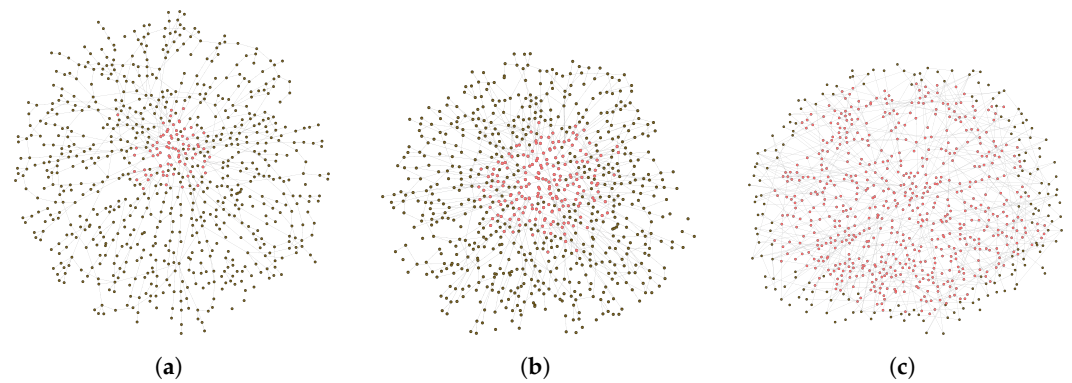


Figure 14. Visualization of subgraphs found by our model in substation secondary screen cabinet terminal block graphical and physical entities. (a) 1000. (b) 2000. (c) 3000.

4. Conclusions

In this paper, the method of matching terminal block schematic and physical information in substation secondary cabinets based on artificial intelligence is studied. Multi-layer object detection networks, tailored to the characteristics of schematic diagrams and physical entities in substation secondary systems, are designed for precise extraction of information. Network topologies oriented towards both schematic and physical systems are then established based on the Neo4j database. The branch-and-bound method is improved by employing a multi-modular graph convolutional network (MGCN) and deep Q-network (DQN), which are subsequently used to solve the maximum common subgraph (MCS), resulting in rapid matching of schematic-to-physical data. Experimental results demonstrate that the algorithm proposed in this paper significantly outperforms traditional algorithms in terms of information extraction in substations and rapid schematic-to-physical matching.

Reflecting on the constructive feedback received, we acknowledge the importance of not only presenting our findings but also suggesting pathways for future enhancements. In light of this, we propose the exploration of innovative labeling techniques, such as the implementation of QR codes on labels, to streamline the identification and tracking processes of terminal blocks within substations further. This approach could potentially offer a more efficient and error-resistant method for managing and accessing detailed information about each component, thereby enhancing the overall effectiveness of our proposed model. We believe that integrating such technologies could serve as a valuable direction for future research, aiming to bridge the gap between current capabilities and the evolving needs of smart grid infrastructure.

Author Contributions: Conceptualization and methodology, W.C. and Z.C.; investigation and writing—original draft preparation, W.C. and C.W.; writing—review and editing, W.C. and T.L. All authors contributed substantially to the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the State Grid Headquarters Science and Technology Project “Research on Unified and Shared Modelling and Intelligent Application Technology for the Whole Secondary System of Substation” (SGHEDK00JYS2200012).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: Author Congying Wu was employed by the company State Grid Economic and Technological Research Institute Co., Ltd. Author Tiecheng Li was employed by the company Power Science and Research Institute of State Grid Hebei Power Co. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Notations

MGCN	Multi-modular graph convolutional network
DQN	Deep Q-network
MCS	Maximum common subgraph
R-CNN	Region-based convolutional neural network
YOLO	You Only Look Once
YOLOv7	YOLO version 7
RL	Reinforcement learning
GNN	Graph neural network
YOLACT	You Only Look At Coefficients
YOLOv7-MS	YOLOv7 with multi-head self-attention mechanism and the space-to-depth module
YOLOv7-DBAH	YOLOv7 based on differentiable binarization attention head
MHSA	Multi-head self-attention mechanism
SPD	Snowflake point deconvolution
YOLOv7-AO	YOLOv7 area-oriented

References

- Li, S.; Gao, Z.; Zhu, Y.; Wang, T.; Rui, M. Research on designing methods of the secondary system in digital substation. In Proceedings of the 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi'an, China, 25–28 October 2016; pp. 289–293.
- James, J.Q.; Markos, C.; Zhang, S. Long-term urban traffic speed prediction with deep learning on graphs. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 7359–7370.
- Zeng, H.; Zhou, Y.; Guo, Q.; Cai, Z.; Sun, H. A distributed deep reinforcement learning-based approach for fast preventive control considering transient stability constraints. *CSEE J. Power Energy Syst.* **2021**, *9*, 197–208.
- Jalali, S.M.J.; Osório, G.J.; Ahmadian, S.; Lotfi, M.; Campos, V.M.; Shafie-khah, M.; Khosravi, A.; Catalão, J.P. New hybrid deep neural architectural search-based ensemble reinforcement learning strategy for wind power forecasting. *IEEE Trans. Ind. Appl.* **2021**, *58*, 15–27. [\[CrossRef\]](#)
- Zhang, N.; Donahue, J.; Girshick, R.; Darrell, T. Part-based R-CNNs for fine-grained category detection. In *Computer Vision—ECCV 2014, Proceedings of the 13th European Conference, Zurich, Switzerland, 6–12 September 2014*; Springer: Cham, Switzerland, 2014; pp. 834–849.
- Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
- Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems 28, Montreal, QC, Canada, 7–12 December 2015.
- Choi, J.; Chun, D.; Kim, H.; Lee, H.J. Gaussian yolov3: An accurate and fast object detector using localization uncertainty for autonomous driving. In Proceedings of the IEEE/CVF INTERNATIONAL Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 502–511.
- Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 2778–2788.
- Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
- 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, 17–24 June 2023; pp. 7464–7475.
- Bai, Y.; Xu, D.; Sun, Y.; Wang, W. Gsearch: Maximum common subgraph detection via learning to search. In Proceedings of the 38th International Conference on Machine Learning, Online, 18–24 July 2021; pp. 588–598.
- Park, Y.; Reeves, D. Deriving common malware behavior through graph clustering. In Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security, Hong Kong, China, 22–24 March 2011; pp. 497–502.
- Cao, N.; Yang, Z.; Wang, C.; Ren, K.; Lou, W. Privacy-preserving query over encrypted graph-structured data in cloud computing. In Proceedings of the 2011 31st International Conference on Distributed Computing Systems, Minneapolis, MN, USA, 20–24 June 2011; pp. 393–402.
- Li, Y.; Gu, C.; Dullien, T.; Vinyals, O.; Kohli, P. Graph matching networks for learning the similarity of graph structured objects. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 3835–3845.
- Ehrlich, H.C.; Rarey, M. Maximum common subgraph isomorphism algorithms and their applications in molecular science: A review. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 68–79. [\[CrossRef\]](#)
- McCreesh, C.; Prosser, P.; Trimble, J. A partitioning algorithm for maximum common subgraph problems. In Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI'17), Melbourne, Australia, 19–25 August 2017; pp. 712–719.

18. Liu, Y.; Li, C.M.; Jiang, H.; He, K. A learning based branch and bound for maximum common subgraph related problems. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 2392–2399.
19. Wang, R.; Yan, J.; Yang, X. Learning combinatorial embedding networks for deep graph matching. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 3056–3065.
20. Bai, Y.; Xu, D.; Gu, K.; Wu, X.; Marinovic, A.; Ro, C.; Sun, Y.; Wang, W. Neural maximum common subgraph detection with guided subgraph extraction. In Proceedings of the ICLR 2020, Addis Ababa, Ethiopia, 30 April 2020.
21. Bolya, D.; Zhou, C.; Xiao, F.; Lee, Y.J. Yolact: Real-time instance segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9157–9166.
22. Radhika, K.; Rao, N.K.; Shalini, N.; Vani, V.D.; Geetavani, B. Driver Drowsiness Detection Using OpenCV and Machine Learning Techniques. In *Microelectronics, Circuits and Systems: Select Proceedings of Micro*; Springer: Singapore, 2021; pp. 413–422.
23. Feng, L.; Ke, Z.; Wu, N. ModelsKG: A Design and Research on Knowledge Graph of Multimodal Curriculum Based on PaddleOCR and DeepKE. In Proceedings of the 14th International Conference on Advanced Computational Intelligence (ICACI), Wuhan, China, 15–17 July 2022; pp. 186–192.
24. Srinivas, A.; Lin, T.Y.; Parmar, N.; Shlens, J.; Abbeel, P.; Vaswani, A. Bottleneck transformers for visual recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 16519–16529.
25. Xiang, P.; Wen, X.; Liu, Y.S.; Cao, Y.P.; Wan, P.; Zheng, W.; Han, Z. Snowflake point deconvolution for point cloud completion and generation with skip-transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 6320–6338. [[CrossRef](#)] [[PubMed](#)]
26. Cao, W.; Chen, Z.; Deng, X.; Wu, C.; Li, T. An Identification Method for Irregular Components Related to Terminal Blocks in Equipment Cabinet of Power Substation. *Sensors* **2023**, *23*, 7739. [[CrossRef](#)] [[PubMed](#)]
27. Hamilton, W.; Ying, Z.; Leskovec, J. Inductive representation learning on large graphs. In Proceedings of the Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 4–9 December 2017.
28. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. In Proceedings of the 6th International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
29. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In *Computer Vision—ECCV 2016, Proceedings of the 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016*; Springer: Cham, Switzerland, 2016; pp. 21–37.
30. Peyré, G.; Cuturi, M.; Solomon, J. Gromov-wasserstein averaging of kernel and distance matrices. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 2664–2672.

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.