International Journal of Information System Modeling and Design
Volume 17 • Issue 1 • January-December 2026

Research on Recognition of Substation Secondary Screen Cabinet Wiring Diagrams Based on VLMs

Yizihe Lang
https://orcid.org/0009-0002-0619-8460
State Grid Jiangsu Electric Power Co., Ltd., China

Chunchao Chen
https://orcid.org/0009-0003-6011-5475
State Grid Jiangsu Electric Power Co., Ltd., China

Qiancheng Cai
https://orcid.org/0009-0000-0717-8029
State Grid Jiangsu Electric Power Co., Ltd., China

Shuangzhu Tao
https://orcid.org/0009-0007-6600-2999
State Grid Jiangsu Electric Power Co., Ltd., China

Xiao Zhang
https://orcid.org/0009-0003-0457-5282
State Grid Xuzhou Power Supply Company, China

Baoxing Ju
https://orcid.org/0009-0008-9425-6352
State Grid Jiangsu Electric Power Co., Ltd., China

Received: August 14th, 2025 | **Accepted:** January 14th, 2026

ABSTRACT

Existing drawing recognition methods struggle with format dependency and limited interpretability, hindering parsing of inconsistent drawings from different institutes. To tackle these issues, this paper proposes a wiring diagram recognition approach leveraging large visual models (such as vision-language models [VLMs]) guided by prompt engineering, enabling analysis of diverse secondary screen cabinet wiring diagrams. Firstly, the influence of prompt engineering styles on VLMs' performance was investigated. Then, using optimized prompts, different VLMs' reliability and parsing accuracy were evaluated. Finally, comparisons with traditional methods and worklist for future engineering applications were made. Results show that, with effective prompt engineering, all tested models demonstrate acceptable drawing recognition ability, with the highest average accuracy reaching 93.48%. However, line complexity and character blurriness introduce interference. Relying on rule-based understanding reduces the reliance on large-scale data training, rendering the method applicable for substation secondary panel wiring diagrams.

KEYWORDS

Visual Recognition, Vision-Language Model, Prompt Engineering, Substation Digitalization, Drawing Parsing

INTRODUCTION

At present, the wave of digitalization has reshaped the operational logic of industries, which demonstrates notable value of artificial intelligence technology empowerment in key fields (He, 2022; Kaur et al., 2022; M. Li & Rohayati, 2024; Sharma et al., 2022). Driven by both energy transition and digital transformation, the digital transformation of power grids in China is accelerating (Tian et al., 2024; H. Wang et al., 2023). It has been found that the construction of substations has become a weak link in digitalization.

DOI: 10.4018/IJISMD.399758

Currently, during the wiring of secondary panels in substations, operators need to manually input drawing data one by one and print cable labels for wiring, which is time consuming and labor intensive. For example, a 500-kV substation equipped with approximately 120 secondary panels and over 1,600 cables requires 4 to 6 workers to input drawing information and print cable caps, resulting in low efficiency and long construction periods. There is an urgent need for more intelligent methods to reduce the steps in manual identification.

This paper proposes a prompt engineering framework that utilizes large models to recognize wiring drawings. It leverages chain-of-thought (CoT) prompting to guide large models in understanding diagram parsing rules. In this work, the influence of prompt styles on model performance was investigated, and the effectiveness of the proposed method across various large visual models was evaluated, focusing on their ability to recognize different wiring diagrams from different design style. A comprehensive analysis is given in this paper based on experimental data, with accuracy as the primary metric, to assess the overall performance of the approach.

Accordingly, the main contributions of this paper are summarized as follows.

1. A method using large visual models (such as vision-language models [VLMs]) with a tailored prompt engineering framework is proposed, which can automatically interpret wiring diagrams by rule-based understanding of electrical schematics.
2. By applying VLMs' image understanding ability to recognize wiring diagrams of a substation secondary panel, the proposed method may relieve the heavy reliance on large amounts of annotated data. It provides a promising solution to recognize different styles of diagrams with minimal training data.

The remainder of this paper is organized as follows. First are reviews of related work, followed by a section that introduces prompt engineering for diagram parsing principles. Then, evaluation of parsing performance is presented, followed by a section that draws the conclusion.

RELATED WORK

The use of computer vision algorithms (D. Wang et al., 2024) to recognize wiring diagrams is an important means of digitalization during the construction phase. The core of this work lies in target detection, accurate text recognition, and clear identification of connection relationships in the diagrams (Bhanbhro et al., 2022; Nguyen et al., 2021; Xiao et al., 2020).

Existing research has mainly focused on using algorithms, such as OpenCV, for image processing and stitching (Ling & Zhu, 2025; L. Wang et al., 2024), optimizing target detection by integrating machine learning or deep learning methods (Hou et al., 2023; Lin et al., 2025) to locate key primitives, and extracting text information from diagrams with the help of character recognition technology to determine the connection relationships between various primitives (W. Liu, 2023; Shi et al., 2023; M. Wang et al., 2024; Y. Wu et al., 2025).

Zhong et al. (2023) proposed a spatiotemporal embedding encoding method, which uses position information to improve the robustness of low-resolution text recognition. Chu et al. (2023), among others, proposed a dual-layer extraction model for effective electrical information regions in diagrams, solving the problem of text localization in complex backgrounds through iteration of image processing and target detection. Lü (2023) and Qu et al. (2024) both utilized the vector graphics characteristics of computer-aided design (CAD) drawings to extract line segment information between connected component symbols and determine the connection relationships between component symbols. H. Liu et al. (2024) used the discriminant locally linear embedding algorithm to enhance the geometric structure features of data, improving the accuracy of connection relationship recognition in secondary wiring diagrams.