

Explaining River Pollution in Hyperspectral Images Using Prompt Engineering

¹Chung-Chian Hsu and ^{*2}Shih-Mao Liu

¹ Department of Information Management
National Yunlin University of Science and Technology, Yunlin, Taiwan,

² Department of Information Management,
National Yunlin University of Science and Technology, Yunlin, Taiwan

*E-mail: oli871113@gmail.com

ABSTRACT

As river pollution becomes an increasingly severe issue, hyperspectral river imagery is gradually becoming a key tool for quickly assessing pollution levels. These image data can be interpreted using large language models (LLMs) to provide pollution analysis of rivers. However, when image complexity is too high or the image quality is poor, the outputs generated by LLMs may not adequately reflect the pollution status. Additionally, interpretations from LLMs without prompt templates may lack depth and coherence, failing to deeply analyze key pollution features in the images. To address these challenges, this study proposes an improved method that converts image data into numerical data and integrates prompt templates into LLM inputs to generate more accurate and comprehensible pollution analyses. A centralized coordinate system is used to precisely describe the river's flow direction and pollution distribution, enhancing the reliability of interpretation results. The study's findings indicate that this method improves the accuracy, comprehensibility, and overall coherence of hyperspectral river image interpretations, thereby promoting the application of AI technology in the field of environmental monitoring.

Keywords: River Pollution, Large Language Model, Prompt template.

1. INTRODUCTION

The rapid development of artificial intelligence technology, particularly the emergence of LLMs such as ChatGPT, has introduced new possibilities for the analysis and interpretation of data and images. LLM leverage natural language processing techniques to quickly analyze vast amounts of data and images, generating easily comprehensible explanations.

Hyperspectral imaging has gradually become a key tool for environmental scientists to assess pollution levels. The integration of LLM into the analysis of hyperspectral images has significantly facilitated the interpretation of

these data. However, when images are overly complex or of poor quality, the explanatory capabilities of LLM may fall short, impacting their effectiveness in environmental monitoring.

In this context, prompt engineering becomes critically important. Prompt engineering involves designing and optimizing prompts to guide LLM in generating more accurate and relevant explanations [1]. By designing specialized prompt templates, LLM can better understand data features, thereby improving the accuracy and comprehensibility of their interpretations. This technique has already demonstrated its potential in various fields, particularly in processing data and generating explanations.

To address the challenge of interpreting hyperspectral images in river pollution monitoring, this study proposes a method that combines image information extraction, data conversion, prompt engineering, and LLM to automatically generate text explanations for complex environmental data. By optimizing the coordinate system and designing specialized prompt templates, the accuracy and comprehensibility of LLM-generated explanations for river pollution images are enhanced. This approach promotes the application of artificial intelligence technology in environmental monitoring and provides valuable reference points for future research and development.

2. BACKGROUND KNOWLEDGE

2.1. Large Language Model (LLM)

Large Language Models (LLMs) are trained on vast amounts of textual data, including articles, books, and media, to learn language structure, semantics, grammar, and context. This enables them to predict, understand, generate text, and answer questions, significantly impacting many NLP tasks [2], such as natural language generation and conversational capabilities.

LLM show great potential in environmental monitoring, especially in river pollution analysis. They can convert complex hyperspectral images into clear

textual descriptions, aiding scientists and decision-makers in formulating effective environmental strategies. Additionally, LLM can generate detailed contextual explanations, enhancing the accuracy of environmental data interpretation and the credibility of AI-driven decisions.

2.2. Prompt Engineering

Prompt engineering is one of the core technologies for interacting with large language models (LLMs). The importance of this technique has become increasingly prominent as LLM are applied across various fields, whether it's for generating code snippets or providing expert consultation services. The goal of prompt engineering is to enable LLM to achieve specific interaction patterns and outcomes by providing a set of precise instructions [3].

By establishing prompt templates to systematize the design of prompts, prompt engineering offers a reusable solution to address common issues in LLM interactions [4]. These prompt patterns encompass aspects such as input semantics, output customization, error identification, prompt refinement, and interaction control. They can be combined to achieve more complex task objectives [5].

Furthermore, the applications of prompt engineering extend beyond simple text generation. By utilizing prompt patterns, LLM can operate in more complex scenarios, such as automatically generating deployment scripts, creating visual charts, or even simulating specific roles or scenarios. These applications demonstrate the significant potential of prompt engineering in enhancing the functionality of LLM and improving the quality of generated outputs [6].

3. MOTHOD

3.1. Enhancing the accuracy and clarity of explanations generated from image-based datasets

This study focuses on enhancing the interpretability and understandability of image-based datasets, particularly those involving hyperspectral aerial images of river pollution. The method combines image information extraction, data conversion, prompt engineering, and LLMs to automatically generate text explanations for complex environmental data.

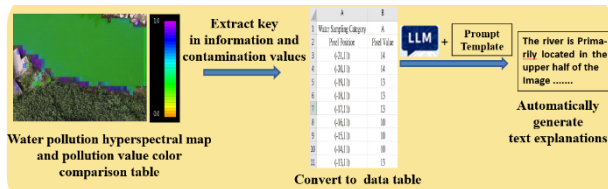


Fig. 1. Flowchart of the explanation generation process for river pollution in hyperspectral aerial images.

3.1.1. Obtaining the corresponding coordinates and relative values of the image

In this study, we first utilized the image processing library (Pillow) in Python to read hyperspectral aerial images of river pollution. To accurately capture the pollution locations and levels within the images, we traversed from the top left to the bottom right corner, extracting the position and corresponding color value of each pixel. These color values were linearly mapped to a pollution value range, where 0 represents no pollution and 1 represents high pollution. The relative pollution values were mapped according to the color gradient shown in Figure 2, which transitions from purple (1.0) to orange (0.0), indicating the progression from high pollution to no pollution.

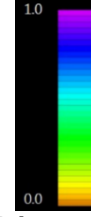


Fig. 2. Color mapping chart

3.1.2. Converting coordinates and pollution values

The first step involves processing the coordinates. Since the coordinates (x, y) are derived by traversing the image from the top left to the bottom right corner, this study uses equations (1) and (2) to transform them into a coordinate system centered at $(0,0)$. This transformation enables the coordinates to extend towards the four corners of the image. In these equations, W and H represent the width and height of the image, respectively.

$$x' = x_i - \frac{W}{2} \quad (1)$$

$$y' = y_i - \frac{H}{2} \quad (2)$$

In this study, Figure 2 was evenly divided into three sections, with each color range corresponding to the Environmental Protection Agency's Biochemical oxygen demand (BOD) standards. These sections represent non-polluted or lightly polluted, moderately polluted, and highly polluted areas, respectively. To further enhance the resolution of the data, the 0-1 range in Figure 2 was converted to a 1-33 scale. In this new scale, values from 1 to 11 represent non-polluted or lightly polluted areas, 12 to 22 represent moderately polluted areas, and 23 to 33 represent highly polluted areas. This finer categorization allows for a more detailed reflection of the actual pollution levels, thereby improving the interpretability and practical value of the data. This transformation ensures that the model can perform more precise quantitative assessments of pollution levels, providing a more accurate data foundation for subsequent analysis.

3.1.3. Generating explanations through language models and partial tables

Since LLMs are primarily designed for text processing and do not support tabular data, CSV files are converted to JSON format upon reading. This JSON data is then input into the LLM for interpretation, enabling the LLM to accurately understand and respond to the information contained in the hyperspectral aerial images of river pollution.

3.1.4. Designing a prompt template for explaining points of interest in images

To enhance the LLM's capabilities in image analysis and interpretation, this study developed a prompt template specifically for interpreting aerial images of river pollution. The design process of this template was carried out in a clear and systematic manner, as illustrated in Fig 3.

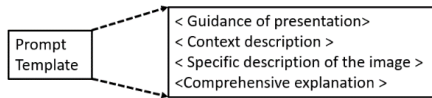


Fig. 3. Structure of the prompt template

1) Guidance of presentation

The first step in designing the prompt template is to establish the overall guidance for interpretation, which includes defining the model's role and objectives. In this study, the model's role is set as an "Environmental Science and Water Quality Assessment Expert," with the objective of analyzing the provided aerial images to assess river pollution levels. This role definition ensures that the model focuses on descriptions related to river pollution when generating content, avoiding the inclusion of irrelevant information.

2) Context description

Detailed settings for the image's background and reference framework were established, including the dimensions of the aerial image, the coordinate system, and its definitions. These settings assist the model in accurately interpreting the spatial information within the image.

3) Specific description of the image

The template guides the model to describe the relative position of the river in the image and analyze the pollution levels, using proportions and sizes to depict the distribution of pollution. The goal is to generate specific and precise descriptions, ensuring that the explanations are easy to understand. By providing examples, the model has a reference point for its responses, which helps ensure consistency and adherence to the intended guidelines.

4) Comprehensive explanation

The template guides the model to generate a comprehensive explanatory report, which includes a summary of the overall pollution levels in the river, an analysis of the variations in specific areas, and the final conclusion provided by the water quality assessment expert. This part of the design ensures that the generated content is coherent and complete, providing a thorough evaluation of the results.

4. EXPERIMENTAL RESULTS

4.1. Direct input and interpretation of water pollution images

In this study, we conducted experiments using the gpt-4o-mini model, which is capable of reading and interpreting images. This model was employed to evaluate the appropriateness of the interpretations. Given that the model's ability in text recognition is already relatively advanced, this study focused primarily on the model's ability to recognize and interpret colors within the images.

Table 1. Direct image input for interpretation

User Prompt
These two pictures are aerial pictures above the river and a comparison table of color pollution values. Aerial photo above the river, the blurred part is the result of hyperspectral pollution labeling as the river flows. The color pollution value comparison table ranges from orange (0) to purple (1) from low pollution to high pollution. Help me analyze and explain 1. Describe the relative position of the river in this aerial photo, 2. Pollution status and location.

Input


Fig 4. River pollution map and color pollution value mapping table

Response by ChatGPT
1. Relative Position of the River: In this aerial image, <u>the river is located in the central lower part of the image, contrasting with the trees and other vegetation areas above it.</u> The edges of the river have a clear boundary with the surrounding grassland, and the flow direction appears to extend from left to right.
2. Pollution Status and Location: Based on the color information in the image, the overall water area appears green, which may indicate higher algae growth or other water quality issues. <u>Specific areas, such as the lower left and right edges, show deeper color changes,</u> like purple and blue, which could represent heavier pollution or changes in water quality.
3. Overall Conclusion: Based on the above analysis, <u>it can be concluded that the overall water quality is poor,</u> indicating a possible phenomenon of eutrophication.

When analyzing this aerial image, several errors were identified in the descriptions. To facilitate illustration, these errors were marked with red underlines in this paper. First, the river is actually located in the upper part of the image, not in the lower central area. This error may have occurred because the color of the river above closely matches the color of the vegetation below, leading to a

misjudgment in positioning. Additionally, the description of color changes along the right edge is also incorrect; in reality, there are no significant color changes along the right edge. Regarding the conclusion about poor overall water quality, relying solely on the color changes in the image is insufficient to reach such a conclusion.

These issues demonstrate that relying solely on language models to automatically interpret images, especially when the image quality is poor, can lead to inaccurate results. Therefore, this study further developed prompt templates to assist the model in interpretation after extracting image information. By providing clear guidance, the model can more accurately describe the river's location, pollution status, and distribution. This approach not only enhances the detail and accuracy of the generated content but also improves the model's performance in handling complex environmental data, making it more effective in image analysis and interpretation.

4.2. Designing prompt templates for interpretation after image information extraction

Due to the complexity of the transformed information, this study designed a detailed prompt template (as shown in Table 2). First, ChatGPT is established as an "Environmental Science and Water Quality Assessment Expert" to enhance the relevance and professionalism of the analysis. Next, the template introduces the image background and coordinate system, adopting a reference framework centered on the midpoint to handle spatial information in the images. Following this, the model describes the relative position of the river and the pollution distribution, ensuring the explanations are specific and easy to understand. Finally, the template guides the model in generating a comprehensive report, covering pollution summaries, regional analysis, and expert evaluations, providing valuable references for environmental research and water quality management.

Table 2. Prompt template

Step 1
1. Role: Environmental science and water quality assessment expert
2. Objective: Analyze the provided aerial image to assess the river pollution situation.
Step 2
Scenario:
1. Aerial image dimensions and range: The photo has dimensions of 42x24. The actual size of the aerial image is 2720x1530, with an aspect ratio of approximately 16:9.
2. Water pollution data:
1. In the json_river_area, the first field represents the location of river pollution, and the second field indicates the pollution value range (1~33), with higher numbers representing more severe pollution.
3. River location description requirements:
(1) When describing relative positions, use descriptions relative to the center of the aerial image, avoiding any numbers or references to rows and columns. The description should be easily understandable.

(2) Describe where the river starts and where it ends within the image.

Step3

Rules:

1. Reference points: Use the center, corners, or edges of the aerial image as reference points to describe the river's location. For example: "The river is located near the upper left corner of the aerial image."

2. Directional terms: Use basic directional terms such as "upper," "lower," "left," "right" to describe the river's location. For example: "The river is in the upper right part of the aerial image, close to the top edge," or "The river is in the lower right portion of the image, near the bottom right corner," or "The river flows along the left edge of the image, turning right as it approaches the bottom edge."

3. Proportion and Size: When describing the river's relative size in the aerial image, use expressions like: "The river occupies the lower half of the image, narrowing as it flows." When describing the river's specific location within the image, you can use proportions, such as: "The river is located in the left third of the image and gradually extends downward."

4. Proportionality:

(1) Overall proportion: Describe the proportion of the river in the aerial image. For example: "The main part of the river is in the lower third of the image, extending upward towards the center."

(2) Local proportion: When the river crosses multiple areas, describe the proportions of each part. For example: "The upper portion of the river occupies the upper left quarter of the image, while the lower portion extends to the lower right quarter."

5. River extension: Describe the river's starting and ending points, overall direction, width changes, and any features. For example: "The river starts from the upper left corner of the image, extends through the central part, and ends in the lower right corner."

Steps 4

Comprehensive explanation:

1. Using coordinate transformation techniques, provide a detailed description of the river's relative position in the 42x24 aerial image based on its location in the picture. Combine reference points and directional language to offer a precise textual description. (River Position Block Description)

2. Describe the overall pollution level of the river, using the relative position within the aerial image for reference. Utilize proportion and size to express the pollution distribution. (Overall Pollution Level Description)

3. If certain areas have significantly different pollution levels compared to the overall situation, specifically describe the location and pollution condition of these areas. (Significant Difference Area Description)

4. As a water quality assessment expert, based on the above analysis, provide a comprehensive conclusion. Consider both overall and local pollution conditions, and offer a general evaluation and recommendations for water quality. (Expert Conclusion)

To validate the effectiveness of the designed prompt template in practical applications, this study utilized the template to analyze aerial images of river pollution. The following experiments demonstrated the application of the template and analyzed the generated explanations, as shown in Table 3.

In the experiments, the underlined sections highlighted the advantages of using numerical analysis over image-based text generation, particularly in the more accurate and clear descriptions of the river's location and pollution status. With precise location information, readers can clearly understand the river's trajectory and pollution distribution, enabling them to effectively identify pollution hotspots and implement targeted remediation measures. This detailed data analysis enhances the accuracy of water quality assessments, increasing the credibility and practicality of the reports, and providing a reliable foundation for environmental protection and management. This study significantly improved the professionalism of LLMs in interpreting aerial images of river pollution, strongly supporting the work in environmental science and water quality assessment.

Table 3. Extracted image data and interpreting using prompt templates.

System Prompt
Using the prompt template (Table 2)
User Prompt
I have a 42*24 pixel picture: X
Input
Extracted coordinate and value information (Figure 4)
Response by ChatGPT
1. Relative position in the aerial image: <u>The river is primarily located in the upper half of the image, especially within the top two-thirds.</u> This portion of the river extends <u>from near the left edge to the right edge</u>, with some areas in the upper section showing signs of pollution. The entire river appears to run parallel to the upper edge of the image. 2. Overall pollution level in the relative position: The pollution is mainly concentrated in the upper half of the image, <u>particularly near the center and slightly to the left, where pollution values are relatively higher</u>, with several spots showing significantly more pollution than other areas. It is worth noting that while there is still some pollution along both sides of the river in the upper and lower parts of the image, the overall water quality is relatively clean. 3. Conclusion: According to the river's water quality assessment, the pollution is relatively light, and <u>the overall water quality is good</u>. The overall assessment indicates light pollution, but localized pollution should still be addressed.

4.3. Comparison of different coordinate systems under the same prompt template

In this study, to compare the impact of different coordinate systems on the generated results under the same prompt template, Tables 4 and 5 were presented. Table 4 shows the interpretation method using a coordinate system with the top-left corner as the origin, while Table 5 demonstrates the interpretation using a centralized coordinate system with the center as the origin.

Table 4. Explanation with the coordinate origin at the top-left corner

System Prompt
Using the prompt template (Table 2)

User Prompt
The input image size is 42x24, with the coordinate system centered at the top-left corner of the image (0,0), and data is traversed from the top-left corner towards the bottom-right corner. Based on this coordinate system, please help me explain and analyze the data in the image.
Input
Extracted coordinate and value information (Figure 4)
Response by ChatGPT
1. River location description: The river starts at the top-left corner of the aerial image, extending to the right edge, covering approximately the upper third of the image. The river flows along the upper edge of the image and gradually extends to the lower-middle section, <u>eventually reaching near the bottom-right corner.</u> 2. River pollution situation: Overall, the river shows slightly elevated pollution levels in the image, <u>with higher pollution values primarily concentrated along the upper edge. The upper section of the river displays indicators of heavy pollution</u>, requiring close monitoring and more detailed water quality assessment and remediation. 3. Overall conclusion: Although the overall water quality assessment of the river indicates light pollution, <u>the concentration of high pollution values in the upper section suggests</u> that enhanced water quality monitoring and remediation measures should be focused on this area.

In the experiments, the underlined sections highlighted errors in the generated descriptions when using the coordinate system with the top-left corner as the origin.

First, in the description of the river's flow direction, it was mentioned that the river starts from the top-left corner and ends near the bottom-right area. However, in reality, the river does not extend to the bottom-right corner but rather terminates in the mid-lower part of the right side. This error likely occurred because the center point was not used as a reference, leading to inaccurate judgments about the actual flow direction and location of the river.

Secondly, the description indicated that high pollution values were concentrated along the upper edge of the river, even suggesting the presence of severe pollution indicators in the upper area. In reality, the high pollution values are concentrated in the mid-lower part of the river, not the upper area. This indicates that without using a centralized coordinate system, the model made errors in judging pollution distribution, mistakenly attributing the mid-lower pollution to the upper region.

To address these issues, the study adopted a centralized coordinate system, with the center of the image as the origin, expecting that this more symmetrical reference framework would improve the accuracy and consistency of the model's generated results. The experimental results are shown in Table 5.

In the experiments, the underlined sections highlighted the advantages of using a centralized coordinate system compared to the top-left corner coordinate system. The descriptions generated under the centralized system more accurately reflected the actual flow direction and location of the river. The description

mentioned that the river generally flows along the left side, slanting to the right, starting from the upper left of the aerial image, extending downward to a mid-lower position, and finally reaching the right side. This is the result of more precise spatial positioning provided by the centralized system.

Furthermore, the centralized system also helps in more clearly identifying pollution sources. The description pointed out that the major pollution sources were concentrated in the area slightly left of the center, which could easily be overlooked or mislocated when using the top-left corner as the origin.

Therefore, it can be concluded that after adopting the centralized coordinate system, the accuracy and consistency of the model's generated descriptions were significantly improved, making it more suitable as the standard coordinate system for future research and applications.

Table 5. Explanation with the coordinate origin centralized

System Prompt
Using the prompt template (Table 2)
User Prompt
The input image size is 42x24, with the coordinate system centered at the image's origin (0,0), and data is traversed outward from the center. Based on this centralized coordinate system, please help me explain and analyze the data in the image. (Extract coordinate and value information from Figure 4)
Input
Extracted coordinate and value information (Figure 4)
Response by ChatGPT
1. The river is located in the upper half of the aerial image, <u>flowing primarily along the left side with a slight curve towards the right. It starts from the upper left corner of the image, extending downward to a point slightly below the center, and finally reaching the right side of the image. The river occupies roughly the upper half of the image, showing a gentle curve throughout its course.</u> 2. Specifically, the river exhibits a noticeable curved flow, with more confined sections concentrated on the left side, <u>gradually slanting to the right as it progresses downward.</u> The pollution levels in some areas, particularly near the center but slightly to the left, indicate moderate pollution. In this section, there are several spots where pollution levels are higher, with these <u>key pollution sources mainly concentrated slightly left of the center point.</u> 3. The overall assessment indicates that, although most areas of the river show low to moderate pollution levels, <u>attention should be given to the hotspots of heavy pollution, especially in the lower-middle section near the left edge.</u> It is recommended to conduct water quality testing and monitoring in these high-pollution hotspots and consider implementing water quality improvement measures to mitigate the impact of the pollution.

5. CONCLUSION

This study proposes an improved method to address the limitations of large language models (LLMs) in interpreting complex or low-quality hyperspectral images for river pollution monitoring. The experiments

conducted included comparing the effects of directly inputting images versus numerical data input into LLMs, analyzing the impact of prompt templates on pollution analysis, and examining the accuracy of different coordinate systems in generating descriptions of river flow direction and pollution distribution. The results indicate that converting images into numerical data combined with prompt templates significantly improves the accuracy and comprehensibility of the interpretations, generating more structured responses. A coordinate system centered at the midpoint proved to be more precise in describing river characteristics compared to the traditional top-left-origin system. This validates the effectiveness of numerical data and centralized coordinate systems, greatly enhancing the performance of LLMs in interpreting river pollution images and providing more reliable analytical results. Future research could build on this by further optimizing prompt templates, exploring other data transformation methods, and extending LLM applications to more environmental issues with localized adjustments to address more complex challenges.

ACKNOWLEDGMENT

This research is supported by the National Science and Technology Council, Taiwan under the project Design of a Command and Control Center for Sustainable Smart Cities (NSTC 112-2634-F-194-001).

REFERENCES

- [1] Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv.
- [2] Phokela, K. K., Sikand, S., Singi, K., Dey, K., Sharma, V. S., & Kaulgud, V. (2023). Smart Prompt Advisor: Multi-objective prompt framework for consistency and best practices. In 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE) (pp. 1-10).
- [3] Spasić, A. J., & Janković, D. S. (2023). Using ChatGPT standard prompt engineering techniques in lesson preparation: Role, instructions and seed-word prompts. In 2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies.
- [4] Naimi, L., Bouziane, E. M., Manaouch, M., & Jakimi, A. (2024). A new approach for automatic test case generation from use case diagram using LLMs and prompt engineering. In 2024 International Conference on Circuit, Systems and Communication (ICCSC) (pp. 1-6). IEEE.
- [5] Rodriguez, A. D., Dearstyne, K. R., & Cleland-Huang, J. (2023). Prompts matter: Insights and strategies for prompt engineering in automated software traceability. In 2023 IEEE 31st International Requirements Engineering Conference Workshops (REW) (pp. 1-6). IEEE.
- [6] Carlander, D., Okada, K., Engström, H., & Kurabayashi, S. (2024). Controlled chain of thought: Eliciting role-play understanding in LLM through prompts. In 2024 IEEE Conference on Games (CoG) (pp. 1-8). IEEE.