Intelligent Satellite Task Planning via Dual-Stage Image Quality Prediction

Wei Peng[1], Jingjie Liu[1], Ye Zhang[1], Lu Li[1], Wei Tan[1], Yimin Liu[1,*]

[1]Research Center for Space Computing System, Zhejiang Laboratory
Hangzhou 311100, China
*liuym@zhejianglab.org

Abstract

Empowered by a data-driven dual-stage prediction framework, this work presents an intelligent satellite task planning–oriented imaging technique that enables high-quality satellite observations. The major challenge in satellite mission planning lies in the inability to forecast image quality during the planning phase. To address this, a comprehensive dataset comprising satellite images and their corresponding weather conditions at acquisition time was constructed. Leveraging this dataset, a dual-stage image quality prediction model was developed to assess observation feasibility with high precision. Experimental results across multiple regions demonstrate a prediction accuracy of approximately 93%, confirming the model's effectiveness and strong potential to enhance mission reliability, enable intelligent satellite task planning, optimize resource utilization, and advance next-generation Earth observation.

Introduction

Satellite remote sensing data play a crucial role in Earth observation, environmental monitoring, resource surveying, and disaster early warning due to their wide spatial coverage, high temporal resolution, and objective observation capabilities (Zhao et al. 2022). However, the quality of satellite imagery is highly affected by various environmental factors such as weather conditions, illumination, and atmospheric disturbances. These uncertainties present significant challenges to imaging mission planning and data acquisition (Li et al. 2022).

In the conventional imaging mode, satellites transmit the acquired data to ground stations for decoding, calibration, and analysis. This "space-sensing, ground-computing" workflow is constrained by the distribution of ground stations, communication bandwidth, and processing capacity (Thangavel et al. 2024). Due to these limitations, less than one-tenth of the captured images can be successfully transmitted. Moreover, even the transmitted data are often degraded by clouds, precipitation, or poor visibility, resulting in reduced usability and wasted observation resources.

Accurately predicting image quality before data acquisition is therefore of great significance for enhancing satellite resource utilization, minimizing ineffective data transmission, and improving the efficiency of satellite task planning (Li et al. 2025). Pre-acquisition quality assessment not only helps select optimal spatiotemporal imaging conditions but also provides critical guidance for intelligent mission scheduling. This capability is particularly valuable for time-sensitive tasks such as disaster monitoring and emergency response, where intelligent satellite task planning based on predicted image quality can substantially improve mission reliability and responsiveness.

However, imaging quality is influenced by multiple interdependent factors, and traditional models often fail to capture these complex nonlinear relationships (Paszkuta 2022). Existing data-driven approaches typically rely on limited high-quality image samples, leading to imbalanced data distributions and weak model generalization.

To overcome these limitations, this paper proposes a data-driven dual-stage image quality prediction model. In contrast to existing methods trained only on high-quality samples or neglecting imaging uncertainties, our model combines classification and regression within a unified framework, effectively handling data imbalance and complex nonlinear relationships. Experiments on GaoFen-2 (GF-2) and GaoFen-7 (GF-7) satellites validate its superior accuracy and robustness for intelligent satellite task planning.

Methodology

Dataset

A total of 22,811 GaoFen-2 (GF-2) and 722 GaoFen-7 (GF-7) imagery scenes were collected from multiple satellites of the national remote sensing platform. The dataset covers all four seasons, and the spatial distribution was selected to maximize coverage across China's geographical regions. For each scene, the acquisition time was extracted from the metadata, and the image center coordinates were computed from the four corner points. Historical meteorological information corresponding to each acquisition was retrieved via the Open-Meteo historical weather API, which provides weather data from 1979 to the present for both domestic and international locations based on latitude, longitude, and date. Retrieved parameters included weather conditions, temperature, precipitation, atmospheric pressure, cloud cover, surface radiation, wind speed and direction, daylength, and sunshine duration. Each image was then paired with its corre-

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

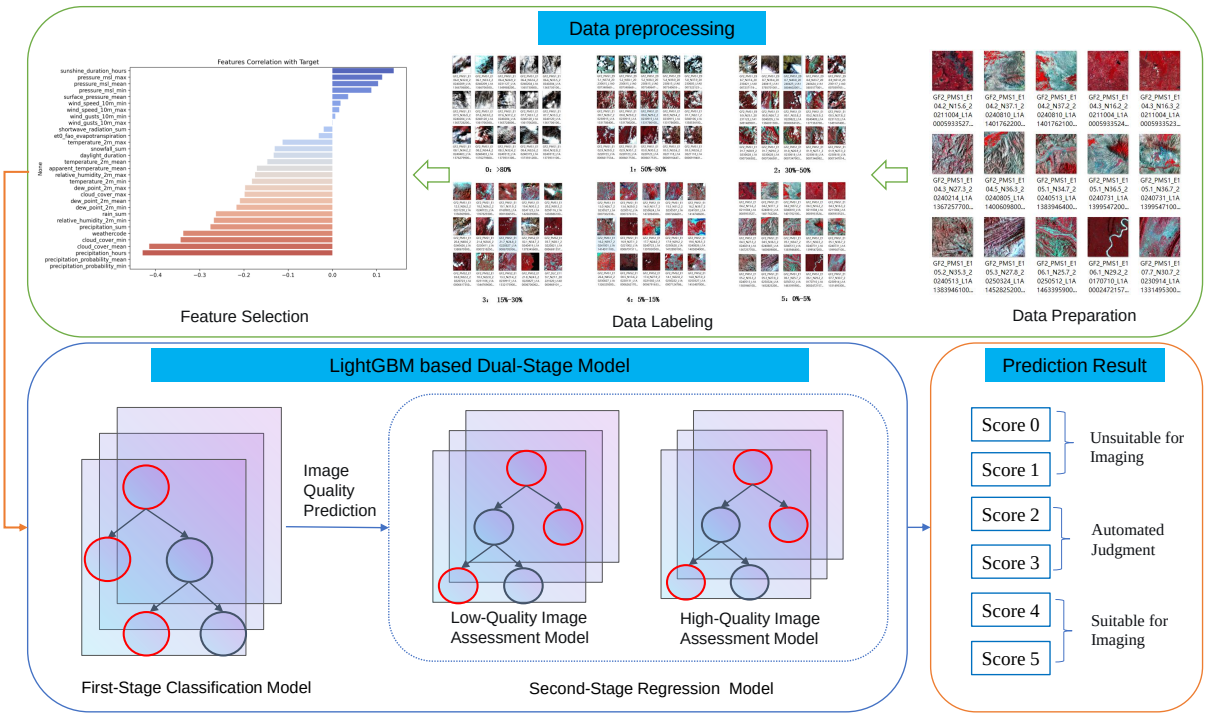

Figure 1: System structure of end-to-end model.

sponding meteorological data to enable integrated analysis.

Data Labeling and Analysis

Data labeling plays a vital role in ensuring the scientific rigor of the experiments and the reliability of their results. The dataset is annotated according to quantitative standards that classify imagery quality into graded levels, as detailed in Table 1.

Table 1: Table of quantitative grading standards.

Score	Cloud Percentage	Description
5	0–5	Nearly cloud-free, optimal
4	5–15	Few clouds, acceptable
3	15–30	Partial clouds, moderate
2	30–50	Many clouds, quality affected
1	50–80	Widespread clouds, poor quality
0	80–100	Almost fully clouded, unusable

We employed a semi-automated approach for labeling the imagery. Initially, a subset of high-quality and low-quality images was manually annotated. The dataset was then quantified and graded using a Cauchy membership function (Viaña et al. 2021), with the primary objective of accurately labeling samples that were either highly suitable or highly unsuitable for imaging—corresponding to scores of 5 and 0, respectively.

It is worth noting that the final labeled dataset exhibits a pronounced class imbalance (Albattah and Khan 2025). Specifically, it contains 758 samples scored as 0, 713 as 1, 348 as 2, 405 as 3, 1,044 as 4, and 19,543 as 5. Images with scores of 4 or higher account for more than 90% of the total dataset.

To enhance the performance of the subsequent predictive model, a preliminary feature analysis was conducted to identify effective weather-related variables. The analysis followed two guiding principles: correlation analysis and redundancy reduction. In the first stage, Pearson correlation coefficients were computed between each weather feature and the target quality score. Features with weak correlations (absolute Pearson coefficient smaller than 0.05) were discarded. The Pearson correlation coefficient is defined as:

$$r_{XY} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \quad (1)$$

where r_{XY} is the Pearson correlation coefficient between variables X and Y, n is the number of paired samples, x_i and y_i are the i-th observations of X and Y, respectively, $\bar{x}$ and $\bar{y}$ denote the mean values of X and Y.

In the second stage, multicollinearity detection was applied to the remaining features by computing their correlation matrix. Among highly correlated feature pairs, only the variable with the stronger correlation to the target was retained. High inter-feature correlation was defined as a coefficient greater than 0.9.

After feature selection, ten representative variables were retained: weather code, average cloud cover, sunshine duration, average sea-level pressure, precipitation amount, precipitation duration, average temperature, average dew point

temperature, average humidity, and day length. Within the dual-stage prediction framework, the importance of these features showed distinct differences. For low-quality prediction, the most influential variables were sea-level pressure, day length, temperature, dew point temperature, sunshine duration, and cloud cover. For high-quality prediction, the ranking shifted to cloud cover, dew point temperature, temperature, sunshine duration, day length, and sea-level pressure.

Dual-Stage Model Based on LightGBM

The proposed model employs LightGBM, a distributed gradient boosting framework based on decision tree algorithms (Ke et al. 2017). The "Light" designation reflects its efficiency in three aspects: smaller sample usage, reduced feature dimensionality, and lower memory consumption. First, LightGBM adopts a histogram-based decision tree construction algorithm, which significantly reduces both the time and space complexity of model training. Second, it utilizes the Gradient-based One-Side Sampling (GOSS) technique, which estimates information gain using only samples with large gradient magnitudes, thereby improving computational efficiency. Most importantly, LightGBM employs a level-wise tree growth strategy that constrains the maximum leaf depth, reducing the risk of overfitting. At each iteration, the leaf with the highest gain is selected for splitting, which decreases computational complexity and enhances model performance under the same number of splits. The overall structure of the model is illustrated in Fig. 2.

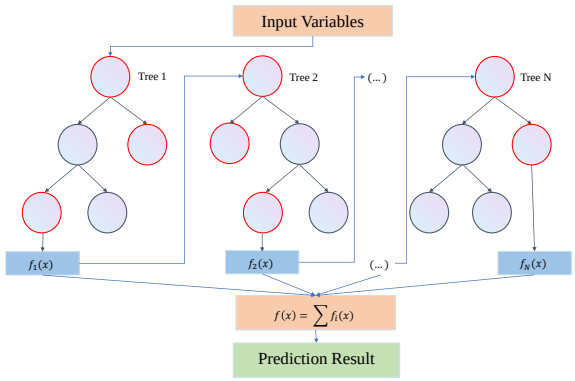

Figure 2: The structure of LightGBM.

Due to the strong correlation between remote sensing image quality and the weather conditions at the time of acquisition, most existing studies only preserve high-quality image samples. To overcome this limitation, we deliberately collected images covering a wide range of quality levels. The resulting dataset, with quality scores ranging from 0 to 5, exhibits a pronounced class imbalance problem.

To achieve accurate prediction of image quality scores, we designed a dual-stage prediction model. In the first stage, a LightGBM-based classification model is employed to distinguish low-quality images from high-quality ones. Specifically, samples with score $\in [0, 3]$ are categorized as low quality, while those with score $\in [4, 5]$ are categorized as high quality. This stage primarily addresses the class imbalance issue by separating the two groups before regression.

In the second stage, regression models are applied within each subset predicted in the first stage. A low-quality regression model is trained for low-quality samples, and a high-quality regression model is trained for high-quality samples. This design enables finer-grained quality estimation while mitigating the effects of data imbalance.

To fully leverage the imbalanced dataset, we optimized the loss function during model training. Both stages are trained jointly under a unified end-to-end optimization framework to obtain optimal parameters. The classification loss in the first stage is defined by the cross-entropy function, while the regression loss in the second stage is measured by the mean squared error (MSE). The overall loss is formulated as a weighted combination of the two components.

The classification loss L_c is expressed as:

$$L_c = \frac{1}{N}\sum_{i=1}^{N}\big[-y_i\log(p_i) - (1-y_i)\log(1-p_i)\big], \quad (2)$$

where y_i denotes the true label of the i-th sample ($y_i = 1$ for high-quality and $y_i = 0$ for low-quality), and p_i represents the predicted probability that sample i belongs to the high-quality class.

The regression loss L_r is defined as:

$$L_r = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2, \quad (3)$$

where y_i and $\hat{y}_i$ represent the ground-truth and predicted quality scores, respectively.

The total joint loss for the two-stage model is then formulated as:

$$L = \lambda L_c + (1-\lambda)L_r, \quad (4)$$

where λ is a weighting coefficient set to 0.5 to balance the classification and regression objectives.

Finally, the complete end-to-end pipeline of the proposed approach—including the classification and regression stages as well as the joint optimization mechanism—is illustrated in Fig.1, providing an intuitive overview of the model architecture.

Experimental Result

Feature analysis conducted during the experimentation yielded the following conclusions. For the first-stage classification model, which distinguished between high and low imaging quality, the determinant features were identified as cloud cover, dew point temperature, ambient temperature, sunshine duration, day length, and atmospheric pressure. In the second-stage regression model, designed for the prediction of imaging quality scores, the critical features were atmospheric pressure, day length, temperature, dew point temperature, sunshine duration, and cloud cover. The relative importance of these features is presented in Fig. 3.

To comprehensively evaluate the proposed model, the Root Mean Square Error (RMSE) and coefficient of determination (R^2) were adopted as key performance metrics. As

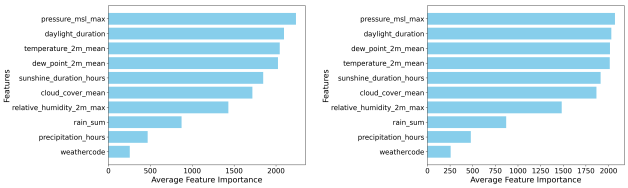

Figure 3: Feature analysis. (a) Feature importance ranking of the classification model in the first stage. (b) Feature importance ranking of the regression model in the second stage.

shown in Fig. 4, the data-driven dual-stage model achieved the lowest prediction error (RMSE = 0.651) and the highest goodness of fit ($R^2 = 0.70$), outperforming all compared methods. Additionally, when employing a large language model (Qwen-turbo) with prompt templates integrating meteorological factors (e.g., precipitation, cloud cover, humidity, pressure, and temperature) to assess imaging feasibility, the prediction accuracies for GF2 and GF7 data were 36.52% and 31.02%, respectively. Under adverse weather (e.g., rain or snow), the accuracy dropped below 20%, indicating limited stability for practical use. Overall, these findings demonstrate notable improvements in both accuracy and robustness, validating the effectiveness of the dual-stage framework for modeling complex satellite imaging conditions.

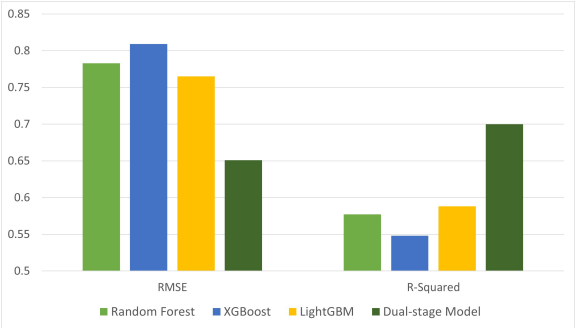

Figure 4: Comparison of performance metrics across different models.

Furthermore, a month-by-month analysis was conducted to evaluate the model's temporal stability and seasonal adaptability. As shown in Fig. 5, the results display a clear seasonal pattern, following the trend Winter > Autumn > Spring > Summer, consistent with meteorological dynamics. Overall, these findings underscore the reliability and generalization ability of the proposed method under diverse environmental conditions.

Conclusion

This work demonstrates an effective satellite tasking-based imaging technique empowered by a data-driven dual-stage prediction model. The proposed approach achieves an imaging quality prediction accuracy of approximately 93% across multiple regions in China, showing strong potential for reliable satellite mission planning and intelligent Earth observation applications. Moreover, by enabling intelligent satellite task planning through accurate pre-acquisition quality prediction, the proposed scheme effectively reduces low-quality or ineffective imaging tasks, thereby minimizing unnecessary satellite resource consumption and improving overall operational efficiency.

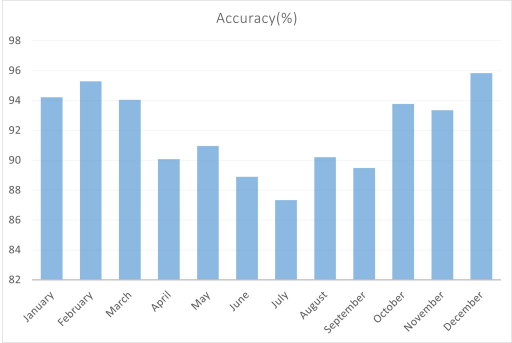

Figure 5: Performance of the proposed model across twelve months.

References

Albattah, W.; and Khan, R. U. 2025. Impact of imbalanced features on large datasets. *Frontiers in Big Data*, 8: 1455442.

Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; and Liu, T.-Y. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.

Li, X.; Chen, Y.; Xing, L.; Chen, Y.; Du, Y.; and He, L. 2025. A review of the frameworks, models, and algorithms for large-scale imaging satellite mission planning. *Expert Systems with Applications*, 292: 128471.

Li, Z.; Zhao, L.; Liu, Y.; Chen, X.; Chen, H.; Zheng, F.; Zhang, Y.; Wang, D.; Li, J.; Liu, J.; et al. 2022. Autonomous mission planning method for optical imaging satellites based on real-time cloud cover information. *Remote Sensing*, 14(11): 2635.

Paszkuta, M. 2022. Impact of cloud cover on local remote sensing–Piaśnica River case study. *Oceanological and Hydrobiological Studies*, 51(3): 283–297.

Thangavel, K.; Sabatini, R.; Gardi, A.; Ranasinghe, K.; Hilton, S.; Servidia, P.; and Spiller, D. 2024. Artificial intelligence for trusted autonomous satellite operations. *Progress in Aerospace Sciences*, 144: 100960.

Viaña, J.; Ralescu, S.; Cohen, K.; Ralescu, A.; and Kreinovich, V. 2021. Why Cauchy membership functions: efficiency.

Zhao, Q.; Yu, L.; Du, Z.; Peng, D.; Hao, P.; Zhang, Y.; and Gong, P. 2022. An overview of the applications of earth observation satellite data: impacts and future trends. *Remote Sensing*, 14(8): 1863.