ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Practical research on UAV Swarm based on SL/T860-2026 specification in long-distance embankment Project Normalized inspection

Lijun Jiang[1,*] *and Lu Chen*[1]

[1] Pujiang River and Lake Management Center, Jinhua, Zhejiang, China

SUMMARY: *Long distance dike inspection requires full section coverage, high-frequency deployment, and reliable defect detection, while manual inspection and stand-alone operation are still limited by efficiency, continuity, and data consistency. Under the SL/T860-2026 framework, a UAV swarm inspection framework integrating multi-source perception, collaborative path planning, multi-source processing, improved detection and specification compliance evaluation is constructed. The system uses six VTOL fixed-wing Uavs, three multi-rotor Uavs, visible light imagery, thermal infrared data, lidar point clouds and synchronous positioning information to support the full section inspection of 72 km of the lede. A three-level cooperative track strategy and track adjustment based on backtracking search are introduced to maintain overlap, edge continuity and safe obstacle avoidance. Field test results show that the coverage integrity, data archipping and specification compliance verification reach 100%, a single 72 km inspection takes 5.2 hours, which is 32 times more efficient than manual inspection, reduces 80% of manpower, and can identify cracks and leakage risks with a width of not less than 1.8mm.*

KEYWORDS: *UAV swarm; Long distance embankment inspection; Multi-source sensor fusion; Cooperative path planning*

1 Introduction

Long distance dike projects undertake tasks such as flood control security, shoreline maintenance and infrastructure protection along the line. The quality of inspection directly affects the stability judgment of the dike, early detection of dangerous situations and the timeliness of disposal. In the scene with large linear span, complex section composition and dense distribution of buildings across the embankment, the inspection system not only needs to complete image acquisition, but also has the computing ability of path organization, spatial positioning, anomaly recognition and result archiving. Relying on the ability of rapid deployment, non-contact observation and multi-load cooperative perception, unmanned aerial vehicle (UAV) platform is becoming an important carrier of intelligent infrastructure inspection.

Puente-Castro et al. [1] reviewed the application of artificial intelligence in UAV swarm path planning, and pointed out that cooperative search, task allocation and dynamic decision-making were the key computing links of UAV swarm operations. Saeed et al. [2] studied the UAV path planning method based on swarm intelligence algorithm, and showed that the path solving efficiency and global convergence ability can be enhanced synchronously under the condition of multiple aircraft. Jeong et al. [3] introduced machine

*750640364@qq.com
https://doi.org/10.65102/is2026994 Published: 30 April, 2026

learning into the bridge inspection process, and proposed a Uavs aided inspection protocol for visibility enhanced images, indicating that visual computing can directly improve the apparent disease recognition effect of infrastructure. Luna et al. [4] proposed a fast multi-UAV path planning method for aerial survey coverage tasks, which made the large-scale airspace coverage have higher spatio-temporal organization efficiency. Espinosa Peralta et al. [5] analyzed the performance of positioning algorithms in two-dimensional and three-dimensional unstructured environments, and showed that the accuracy of laser sensing and spatial registration determined the geometric reliability of inspection results.

Flores Pena et al. [6] proposed a multi-UAV remote sensing operation system, indicating that platform cooperation, link communication and data consistency control are the basic support in engineering deployment. Shafiq et al. [7] studied the convergence characteristics of multi-UAV track based on max-min ant colony optimization, which provided the algorithm basis for group route generation under complex constraints. Avdelidis et al. [8] constructed a defect recognition algorithm around the UAV visual inspection, and promoted the apparent disease extraction to be driven by the model instead of manual interpretation. Gugan et al. [9] summarized the development direction of autonomous UAV path planning, emphasizing real-time re-planning, risk aversion and edge response capabilities in complex environments. Cetinsaya et al. [10] reviewed the development process from traditional control to cluster collaboration, and believed that multi-computer control, intelligent perception and task autonomy were forming a unified technology chain.

The normal inspection of embankments is different from the general scene, and the operation objects show a significant linear distribution feature. The upwater slope, backwater slope, embankment roof, berm ground and dike crossing buildings together form a continuous and heterogeneous inspection space. Single machine operation can complete local acquisition, but it is easy to be constrained by endurance, edge consistency and transition efficiency in 100-km tasks, and manual inspection is also difficult to support high-frequency, full-section and standardized output. After the implementation of SL/T860-2026, the inspection activities have formed clearer technical constraints in terms of spatial accuracy, coverage integrity, frequency arrangement, result format, safety management and quality acceptance, and the embankment inspection has been transformed into a computing system construction process oriented to specification execution.

Based on this background, this paper focuses on the requirements of long-distance embankment inspection under the constraints of specifications, constructs a normalized inspection framework of UAV swarm, which integrates multi-source sensor information, improved detection model, backtracking search cooperative path planning and specification compliance evaluation, and combines the engineering adaptation ability of the 72 km embankment operation verification system. Section 2 reviews related research, Section 3 describes the methodology, Section 4 presents the results, Section 5 discusses, and Section 6 concludes.

2 Related work

The research of UAV inspection for long-distance linear infrastructure has gradually shifted from single-camera shooting and manual interpretation to the computing paradigm of cluster collaboration, visual recognition and autonomous planning fusion in recent years. Katkuri et al. [11] reviewed the research on autonomous navigation based on the vision framework of deep learning, and pointed out that the perception model, path decision and airborne computing power configuration have formed a tight coupling. Ejaz et al. [12] reviewed UAV

image computing in infrastructure management, and argued that crack, surface deformation and component anomaly extraction are shifting from traditional feature engineering to end-to-end vision model, but the stability of small target recognition under complex background is still affected by scene disturbance. Luna et al. [13] proposed a multi-UAV coverage path system with in-flight re-planning capability, which enables the local track correction of large-scale regional inspection in the task execution stage, indicating that multi-aircraft collaboration has moved from offline planning to online adjustment.

In terms of safety constraints and airborne perception, Boukabou et al. [14] analyzed the safety distance between Uavs and the electromagnetic environment during UHV transmission line inspection, which provided a quantitative basis for the operation boundary modeling in high-risk scenarios. Hansen et al. [15] proposed an active target detection and tracking method based on pan-tilt mechanism, so that the UAV can maintain the continuous observation ability of the target area in the dynamic flight state. Santos et al. [16] jointly used visible light and thermal infrared information for power line recognition, and proved that multi-modal vision was helpful to improve the detection reliability in weak texture background. Takaya et al. [17] constructed a quadrotor power transmission inspection system, forming a complete engineering chain from flight control to target acquisition, indicating that platform-level collaborative design has a direct impact on the stable output of inspection results.

Focusing on cluster track organization and task solving, Quadt et al. [18] summarized the research on multi-objective UAV path planning, and pointed out that the range cost, risk aversion, time constraint and coverage benefit often need to be solved simultaneously. Panowicz et al. [19] proposed a robust optimization model for UAV swarm tasks, so that the task planning can maintain the enforceability under disturbance conditions, which provides a method basis for cross-section relay operation in long-distance inspection. Adoni et al. [20] constructed a self-organizing intelligent cluster based on the master-slave paradigm, and completed the verification of autonomous task planning, indicating that the group control structure has been able to support more complex collaborative execution processes. In order to more clearly present the technical path and applicable characteristics of the existing results, Table 1 gives a comparative analysis of related studies.

Table 1: Comparative analysis of studies related to UAV inspection

Reference	Research Focus	Core Method	Limitation	Implication
[11]	Review of autonomous navigation	Deep vision-based perception	Large variation in scene adaptation	Strengthen adaptation to specification-constrained scenarios
[12]	Infrastructure image management	Computer vision-based inspection	Small-target stability is susceptible to disturbance	Enhance the identification of tiny embankment defects
[13]	Multi-UAV coverage planning	In-flight replanning	Insufficient representation of linear engineering constraints	Introduce long-distance collaborative mechanisms
[14]	Inspection boundaries in high-risk environments	Safety distance modeling	Strong scene specificity	Extend safety control for normalized inspection
[15]	Active target tracking	Gimbal-driven detection	Weak capability in organizing inspection outputs	Link inspection result generation with sensing process
[16]	Visible and thermal infrared recognition	Multimodal deep learning	Limited discussion on continuous full-section coverage	Improve consistency of multi-source data fusion
[17]	Quadrotor inspection system	Platform-level system construction	Limited single-UAV operation range	Develop swarm-based collaborative inspection mode
[18]	Review of multi-objective trajectory planning	Multi-objective optimization	High computational complexity	Balance efficiency and accuracy
[19]	Robust swarm planning	Robust optimization model	Insufficient use of real-time image feedback	Combine with online perception updates
[20]	Self-organized swarm planning	Leader–follower collaborative mechanism	Lack of specification-driven validation	Incorporate specification-constrained verification

The existing research has formed a strong accumulation in autonomous navigation, multi-modal recognition, online re-planning and cluster control, but most of the results serve Bridges, power lines, general aerial survey or open area perception, and the research on the normalized inspection of embankment projects is still mainly based on local application. In the scene with long linear distance, many section elements, strict results format and clear frequency requirements, it is difficult to simultaneously ensure track continuity, edge accuracy, hidden danger identification consistency and result compliance output by only relying on general visual inspection or conventional coverage planning. After the implementation of SL/T860-2026, the inspection activities have formed more refined technical boundaries in terms of coverage integrity, data accuracy, result archiving, security control and quality acceptance, and the input organization, collaborative scheduling and result generation of the computational model have also changed. Based on this, this paper puts the specification constraint, cluster collaboration, disease identification and quality control into the same

computing framework, and constructs a normalized inspection method suitable for long-distance embankment projects, so as to realize the whole process closed-loop from task modeling, perception processing to result evaluation.

From the perspective of application objects, the existing inspection research mainly focuses on bridge components, transmission corridors and general remote sensing scenes, and these results provide transferable methods in target detection, geometric positioning and online re-planning. However, embankment inspection has the characteristics of continuous strip distribution, synchronous coverage of water slopes against and behind the embankment, coexistence of buildings across the embankment and hidden dangers, and inspection frequency constrained by industry norms. Therefore, higher requirements are put forward for multi-machine relay, edge consistency, whole segment coordinate unification and standardized output. Especially under the condition of regular inspection, the algorithm not only needs to complete target recognition, but also supports the overall calculation chain of task decomposition, track connection, data fusion and achievement archiving. Existing research has little to do with the regulation-driven inspection process, and there is still a lack of systematic implementation paths for embankment scenarios on how to complete the standardization of multi-source data under a unified coordinate benchmark, how to convert the identification results into traceable inspection accounts, and how to maintain the stability control of navigation height and overlap rate in long-distance tasks. This also makes the related research from the performance comparison of single model further to the research of system collaboration for engineering deployment. Therefore, the cluster inspection method for canonical closed-loop control has practical significance.

3 Methods

3.1 Data acquisition and task modeling of long-distance embankment inspection under the constraints of SL/T860-2026

The data acquisition of long-distance embankment inspection is not a simple flight framing process, but a task modeling process constrained by standard indicators. The research object is a 72 km secondary dike, along which 30 dike piercing buildings and 12 dangerous sections are distributed, and the spatial shape is spread out in a continuous strip. According to the requirements of the regular inspection, the system integrates the embankment roof, the water slope, the back slope and the berm into the sampling domain, and organizes the fixed-wing macro acquisition, the multi-rotor fine mining and the ground terminal scheduling into the same data link. The acquisition terminal is equipped with six VTOL fixed-wing Uavs, three industrial-grade multi-rotor Uavs and a set of cluster ground control terminals, which synchronously access visible light images, infrared thermal imaging, light laser radar, POS positioning and attitude information and flight status log, forming a heterogeneous acquisition system that takes into account both large-scale coverage and detailed inspection of key parts.

To intuitively illustrate the corresponding processes among levee unit segmentation, platform division of labor, data flow organization and scheduling relationships, Fig. 1 shows the data acquisition and task modeling process.

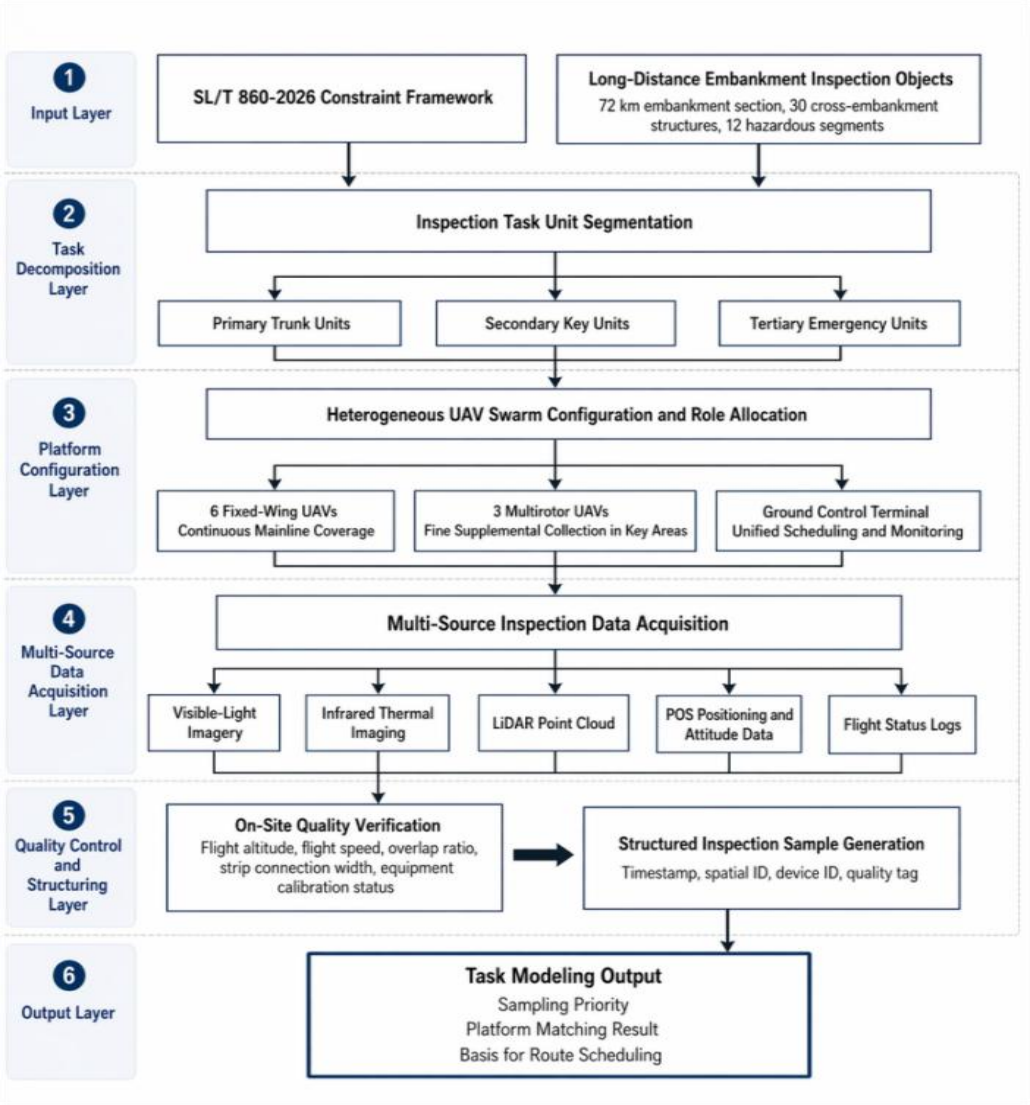

Figure 1: Flowchart of data collection and task modeling for long-distance dike inspection

In order to achieve quantitative expression of long-distance embankment division, heterogeneous platform matching and sampling load distribution under uniform specification constraints, this paper defines the task intensity as follows:

$$S_k=\alpha\frac{L_k}{\bar{L}}+\beta\frac{N_k}{\bar{N}}+\gamma\frac{H_k}{\bar{H}}+\delta\frac{C_k}{\bar{C}} \tag{1}$$

where, S_k represents the task intensity of the k inspection unit, L_k represents the unit length, N_k represents the number of buildings crossing the dike, H_k represents the density of historical hidden dangers, C_k represents the complexity of terrain and occlusion, α, β, γ, δ are the weight coefficients, $\bar{L}$, $\bar{N}$, $\bar{H}$, $\bar{C}$ represent the sample mean of the corresponding variable respectively. The formula maps linear length, structure distribution and risk characteristics into a unified evaluation space, which can be used to determine the cell division boundary, platform placement sequence and sampling priority.

In order to avoid coverage breakage, insufficient overlap and flight time imbalance in long-distance relay operation, this paper further constructs the unit scheduling constraint model as follows:

$$T_k = \max\left(\frac{L_k}{v_f}, \eta_k \frac{L_k}{v_m}\right) + \tau_k + \lambda\Omega_k \tag{2}$$

where, T_k represents the estimated operation time of the k unit, $\frac{L_k}{v_f}$ represents the time required for the fixed wing to complete the acquisition of the main navigation zone, $\eta_k \frac{L_k}{v_m}$ represents the reacquisition time of the multi-rotor in the key area, η_k is the proportion coefficient of the key part, τ_k represents the additional time of take-off and landing, transition and verification, and Ω_k represents the integrity constraint term of the connection edge. This formula is used to uniformly measure the acquisition time and the edge quality, so that the relay unit can complete the sequential scheduling under the condition of satisfying the standard coverage and overlap rate.

On this basis, field operations are written into a unified metadata table according to the process of pre-flight calibration, in-flight monitoring and post-flight verification. A total of 32800 valid visible light images, 186GB thermal infrared data and 124GB point cloud data were obtained from the original acquisition, and all samples were append with time stamps, space numbers, equipment numbers and quality marks. The resulting data is no longer discrete material, but structured input that can be directly fed into subsequent identification, planning, and archiving modules. And maintain the stable comparability and continuous tracking of intertemporal inspection results.

3.2 Embankment multi-source data preprocessing and feature standardization for normalized inspection

The multi-source data of embankments for regular inspection cannot be directly entered into the identification and evaluation model, and it needs to be unified at three levels of time, space and dimension. Visible image, thermal infrared data, laser point cloud, POS trajectory and flying target have obvious differences in sampling frequency, observation range and noise structure. If directly stitching, it is easy to cause boundary drift, weak feature submerging and cross-unit label ambiguity. To this end, the preprocessing stage firstly executes coordinate unification, time alignment and anomaly removal, and then performs resampling, slice coding and risk labeling according to two types of tasks of macro inspection and fine inspection respectively, so that the same embankment section can form aligned, retrievable and traceable structured input under different loads.

In order to maintain strict uniform constraints on the spatial scale, numerical dimension and temporal sampling frequency of multi-source observations, this paper defines the hierarchical normalization expression as follows.

$$Y_{i,t} = \omega_i \frac{x_{i,t} - \mu_i}{\sigma_i + \varepsilon} + (1 - \omega_i) \frac{x_{i,t} - m_i}{M_i - m_i + \varepsilon} \tag{3}$$

Here, $Y_{i,t}$ represents the standardized result of the i type sensing feature at time t, $x_{i,t}$ represents the original observation value, μ_i and σ_i represent the mean and standard deviation, M_i and m_i represent the minimum and maximum value, ω_i represent the modal weight, and ε is the stability constant. This formula puts standard deviation normalization and interval scaling into the same framework, which can preserve global distribution characteristics and local boundary information at the same time, and reduce the suppression of small crack texture, temperature anomaly and local elevation mutation by single normalization method.

To illustrate the connection between coordinate unification, time alignment, quality screening and feature generation for multi-source data, Fig. 2 shows the processing flow of feature standardization.

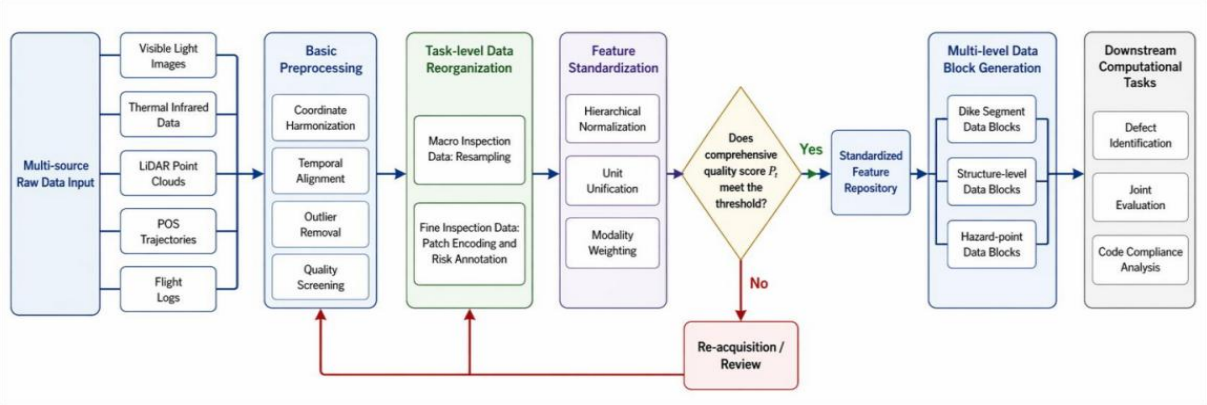

Figure 2: Flow chart of embankment multi-source data preprocessing and feature standardization

In order to make standardized features form a unified input in subsequent identification, edge checking and standard evaluation, this paper further defines the quality aggregation function as follows:

$$P_t=aq_t+br_t+cg_t+\rho d_t \tag{4}$$

where P_t represents the comprehensive quality score at time t; q_t, r_t, g_t, d_t represent image clarity, overlap rate up to scale, geometric registration accuracy, and data integrity, respectively; a, b, c, and ρ are weight parameters. This formula is not a simple score, but is used to screen whether to enter the subsequent identification and achievement production process. When P_t is lower than the threshold, the sample is marked as supplementary collection or review status. When P_t meets the threshold, the sample enters the link of defect identification, edge connection evaluation and ledger generation.

After the above processing, all samples are organized into three types of data blocks: embankment layer, structure layer and hidden danger layer, and each type of data block is also attached with space number, time stamp, platform source and quality label. The feature set formed in this way not only meets the requirements of the specification for consistency, traceability and standardization of results, but also provides a stable calculation basis for subsequent disease identification, track evaluation and compliance analysis. For the cross-cell edge area, the system also performs the same name point check and time window complex comparison to control the boundary drift and label conflict caused by multi-machine relay collection, and improves the spatial consistency of the whole segment samples before entering the model.

3.3 UAV swarm normalized inspection framework for long-distance embankment projects

3.3.1 Improved detection model for dike disease and risk point identification

The target recognition in the regular inspection of long-distance embankments is different from the general aerial scene. The objects include not only fine cracks, gullies, collapses and scour at the foot of the embankment whose width is close to the lower limit of the code

identification, but also hidden dangers of leakage dominated by thermal anomalies and illegal encroachment along the line. Due to the spatial scale difference between the input formed by fixed-wing macro inspection and multi-rotor fine inspection, the single-scale detector is prone to weak response of thin targets, edge fracture and multi-modal inconsistency in long strip scenes. To this end, this paper constructs an improved detection model, which unifies visible light, infrared and local elevation features, and identifies diseases and risk points through three steps: direction-sensitive convolution, cross-scale reorganization and risk-guided classification. The model inputs a uniform slice of 640×640 local inspection block, and retains the original embankment section number and timestamp, so that the identification results can be traced back to the specific section and job batch.

In order to accurately represent the hierarchical convolution mapping relationship of multi-source inspection blocks in the backbone network, this paper constructs the initial feature extraction expression as follows.

$$F^{(l)}=\phi\left(W^{(l)}*F^{(l-1)}+b^{(l)}\right),\quad F^{(0)}=X \tag{5}$$

where X represents the input tensor formed by the concatenation of visible light texture, infrared temperature difference and local elevation channel, $F^{(l)}$ represents the output feature of the l layer, $W^{(l)}$ and $b^{(l)}$ represent the convolution kernel weight and bias term, respectively, * represents the convolution operation, and $\phi(\cdot)$ represents the nonlinear activation function. This equation is used to describe the layer-by-layer mapping process of the input data in the backbone network, and its role is to convert the edges, temperature differences and elevation disturbances in the original image into semantic features that can be used for disease recognition.

In order to avoid the weakening of small-scale cracks in the long strip embankment image during the deep downsampling process, this paper further defines the cross-scale feature reconstruction process as follows.

$$\tilde{F}_s=U(F_{s+1})\oplus C(F_s)\oplus D(F_{s-1}) \tag{6}$$

Here, $\tilde{F}_s$ represents the feature after reorganization at the s layer, $U(\cdot)$ represents the upsampling operator, $C(\cdot)$ represents the current scale convolution enhancement, $D(\cdot)$ represents the downsampling complementary path, and $\oplus$ represents the channle-level concatenation operation. The core of this formula is to merge the texture and semantics of adjacent scales into the current layer, so that mm-scale cracks, local wet patches and small slope collapses maintain continuous responses at different resolutions, and reduce the loss of targets caused by scale switching.

In order to clearly show the internal organization relationship of the improved detection model from multi-source input to risk output, Fig. 3 shows the network structure for dike disease and risk point identification.

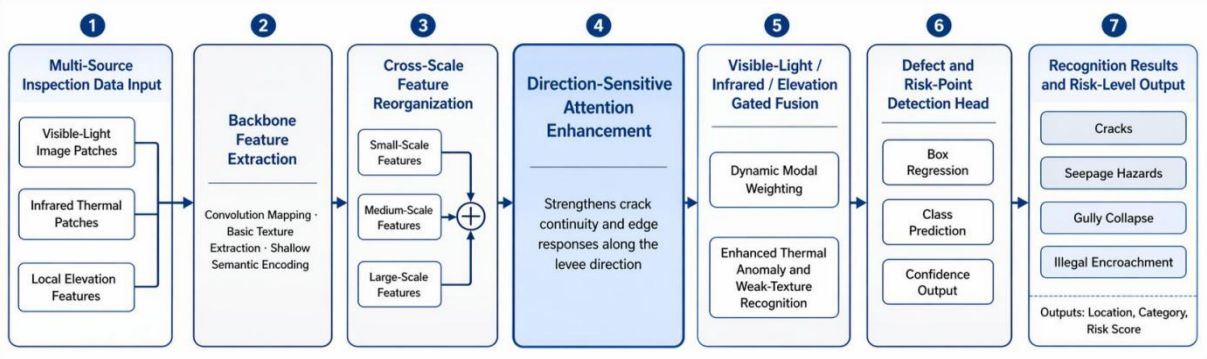

Figure 3: Structural diagram of the improved detection model for dike disease and risk point identification

After the cross-scale reconstruction, the model needs to further strengthen the slender texture along the embankment direction and the lateral failure boundary. To this end, this paper sets the direction sensitive attention mechanism, whose feature weighting process is defined as follows.

$$A_{ij}=\frac{\exp\left(q_i^{\top}k_j+\lambda\cos\theta_{ij}\right)}{\sum_{j=1}^{N}\exp\left(q_i^{\top}k_j+\lambda\cos\theta_{ij}\right)} \tag{7}$$

Here, A_{ij} represents the attention weight of location i to location j, q_i and k_j represent the query vector and key vector respectively, θ_{ij} represents the Angle between the two locations in the main direction of the embankment, and λ represents the direction constraint coefficient. This formula combines the similarity term and the direction consistency term in the traditional attention, so that the model is more inclined to retain the linear cracks extending along the dike and the gully targets with continuous edges, thereby reducing the interference of grass occlusion and shadow fracture on the recognition results.

Considering that covert leakage often manifests as thermal anomalies appearing before apparent texture anomalies, this paper further defines the modal fusion expression of visible and infrared features as follows.

$$Z=\sigma(W_vF_v+W_tF_t+W_hF_h)\odot F_v+\left(1-\sigma(W_vF_v+W_tF_t+W_hF_h)\right)\odot F_t \tag{8}$$

where F_v, F_t and F_h represent visible light, thermal infrared and elevation features respectively, W_v, W_t and W_h represent the corresponding weight matrix, $\sigma(\cdot)$represents the gating function, and $\odot$ represents element-wise multiplication. This formula dynamically assigns weights to different modes by gating. When the apparent texture is clear, the visible light component is dominant. When weak light, wet spots or occlusion weaken the visible light response, the infrared thermal anomaly features obtain higher weight, thereby improving the detection stability of hidden leakage and slope wet zone.

In the target box regression stage, in order to take into account the location deviation, scale deviation and the continuity of the edge of the band disease, this paper constructs the following localization loss function.

$$L_{loc}=1-IoU+\frac{\rho^2(c,c^*)}{d^2}+\alpha\left(\frac{w-w^*}{w^*}\right)^2+\beta\left(\frac{h-h^*}{h^*}\right)^2+\gamma\left(1-\frac{l}{l^*}\right)^2 \tag{9}$$

where IoU represents the intersection ratio between the predicted box and the real box, $\rho^2(c,c^*)$ represents the square distance between the center point, d represents the diagonal length of the minimum bounding box, w, h and w^*,h^* represent the width and height of the predicted box and the real box respectively, l and l^* represent the continuous length of the edge, α, β, β are weight parameters. Compared with the conventional loss that only emphasizes the rectangular overlap degree, the continuous length constraint is additionally introduced, which makes the model more sensitive to slender cracks, strip scour and slope edge failure.

In the final output step, the model not only gives the category and location, but also estimates the level of risk points. To this end, this paper defines the risk score calculation as follows.

$$R_n=\mu_1 p_n+\mu_2 \frac{\Delta T_n}{\Delta T_{max}}+\mu_3 \frac{w_n}{w_{max}}+\mu_4 \frac{l_n}{l_{max}}+\mu_5 g_n \tag{10}$$

Here, R_n represents the risk score of the n candidate target, p_n represents the detection confidence, ΔT_n represents the local temperature difference, w_n represents the width of the crack or abnormal region, l_n represents the target extension length, g_n represents the weight of the structure sensitive area, and μ_1 to μ_5 are the normalized weights. This formula enables the recognition results to directly enter the risk ranking and standardized evaluation process, and no longer stay at the category output level of general target detection.

Through the above design, the improved detection model has completed the unification of multi-source input, cross-scale expression, direction constraint and risk output in the structure. The results obtained in this way can not only support the automatic identification of embankment cracks, leakage risks, gully and illegal encroachment, but also provide an interpretable basis for disease parameters for subsequent specification compliance evaluation and key area review scheduling.

3.3.2 Cluster cooperative path planning and safe obstacle avoidance mechanism based on Backtracking search

There are obvious differences between cluster cooperative path planning and general coverage path planning for long distance embankment inspection. The task not only requires multiple Uavs to achieve continuous relay in a 72 km strip space, but also needs to meet the control requirements of heading overlap rate, lateral overlap rate, edge width and height, while taking into account the key area acquisition, cross-platform task decomposition and operation safety. In this paper, a backtracking search mechanism is introduced in the path planning part to incorporate task allocation, dynamic replanning and obstacle avoidance control into a unified optimization framework. The fixed wing is mainly responsible for the main line inspection, the multi-rotor is responsible for the fine inspection of key areas, and the ground terminal uses the real-time link to update the cluster status, forming a collaborative process of "offline initialization - online correction - conflict avoidance - result feedback".

In order to uniformly represent the comprehensive operation cost of each inspection unit in the stage of platform selection and task allocation, this paper defines the scheduling cost function as follows.

$$J_{m,k}=\alpha_1 \frac{L_k}{V_m}+\alpha_2 E_{m,k}+\alpha_3 B_{m,k}+\alpha_4 U_{m,k} \tag{11}$$

Here, $J_{m,k}$ represents the scheduling cost of platform m executing the k inspection unit, L_k represents the unit length, V_m represents the effective cruise speed of platform, $E_{m,k}$ represents the energy consumption cost, $B_{m,k}$ represents the edge compensation cost, $U_{m,k}$ represents the supplement cost of key areas, and α_1 to α_4 are weight parameters. The function of this equation is to incorporate the range, energy consumption, edge connection accuracy and local acquisition requirements into the same evaluation space, so as to provide a quantitative basis for the task allocation between fixed wing and multi-rotor.

In the track generation process, the environmental risk needs to participate in the update in the form of a continuous field. To this end, this paper constructs the following dynamic threat field expression.

$$\Psi(x,y,t)=\sum_{j=1}^{M}\frac{\kappa_j}{\left\|(x,y)-o_j(t)\right\|^2+\epsilon}+\zeta_1 W(x,y)+\zeta_2 R(x,y) \tag{12}$$

Here, $\Psi(x,y,t)$ represents the comprehensive threat value at the position (x,y) at time t, $o_j(t)$ represents the position of the j dynamic obstacle, κ_j represents the hazard weight of the obstacle, $W(x,y)$ represents the wind field disturbance term, $R(x,y)$ represents the constraint term of the no-fly or high-risk area, ϵ is the stability constant, ζ_1, ζ_2 are the regulation coefficients. This formula not only deals with static structures and temporary obstacles, but also includes wind disturbance and restricted airspace into the risk field at the same time, so that the flight path planning is more in line with the actual safety constraints in the normal inspection process.

In order to easily show the key parameter Settings of the backtracking search mechanism in cluster cooperative path planning, Table 2 shows the main optimization parameters.

Table 2: Key parameter Settings for backtracking search cluster collaborative path planning

Parameter	Meaning	Value
(M)	Population size	40
(T)	Maximum number of iterations	120
(P_c)	Crossover probability	0.75
(P_m)	Mutation probability	0.30
(Δ_h)	Allowable flight altitude deviation	3% of the designed flight altitude
(Δ_d)	Safety distance threshold	20 m
(Ω)	Minimum overlap width at stitching boundary	250 m

In the backtracking search update stage, this paper adopts the track correction form of superposition of historical optimal guidance and random disturbance, and its calculation is expressed as follows.

$$X_i^{t+1}=X_i^t+\lambda_t\left(X_i^t-X_i^{old}\right)+\nu_t\left(X_{best}^t-X_i^t\right)+\xi_t \tag{13}$$

Here, X_i^t represents the state vector of the i candidate track at generation t, X_i^{old} represents the historical archived track, X_{best}^t represents the current optimal track, λ_t represents the backtracking factor, ν_t represents the shrinkage guidance coefficient, and ξ_t represents the random disturbance term. The meaning of this formula is that the stability of historical feasible paths is preserved, and the current optimal result is used to accelerate the convergence, so that the cluster can achieve a better compromise between edge continuity,

obstacle avoidance and total task time.

In order to more clearly illustrate the internal connection relationship between cluster cooperation, risk field update and online obstacle avoidance, Fig. 4 shows the cluster cooperative path planning and safe obstacle avoidance mechanism based on backtracking search.

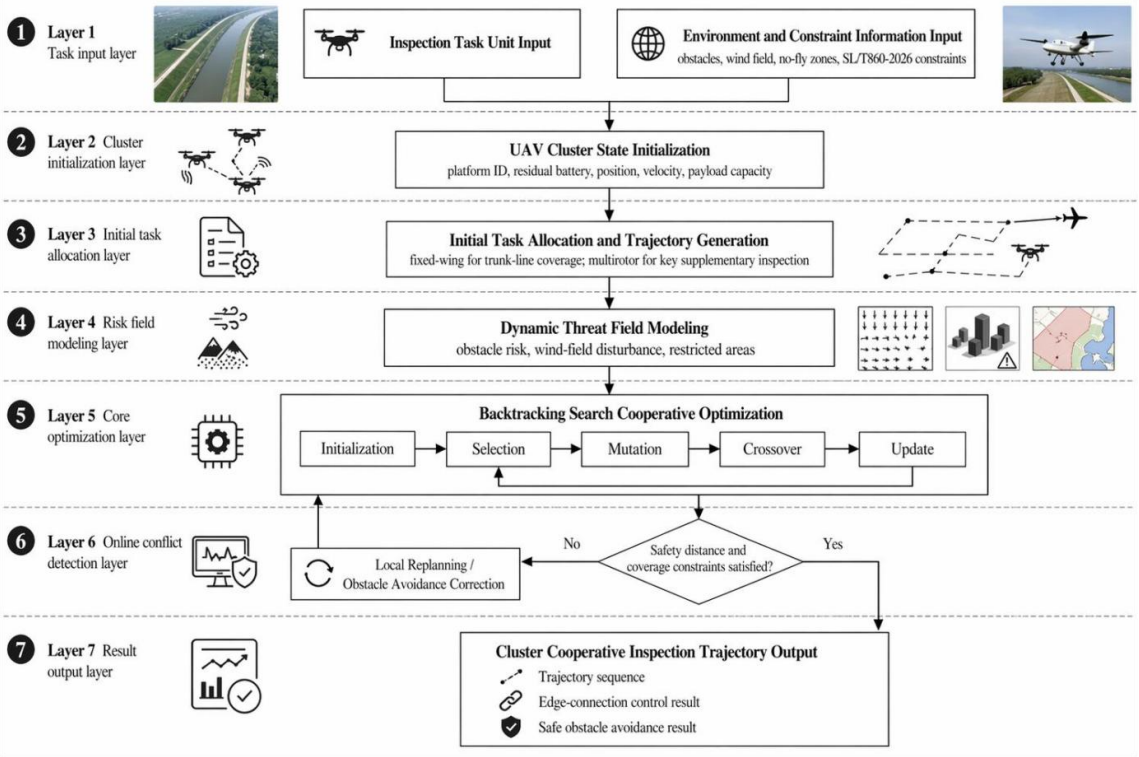

Figure 4: Diagram of cluster cooperative path planning and safe obstacle avoidance mechanism based on backtracking search

Under the condition of multiple machines working at the same time, the requirements of space separation and time peak error must be satisfied between the platforms. To this end, the safe obstacle avoidance constraint between Uavs is defined in this paper as follows.

$$C_{ij}(t)=\left\|p_i(t)-p_j(t)\right\|-\Delta_d-\chi_{ij}(t)\geq 0 \tag{14}$$

Here, $C_{ij}(t)$ represents the safety margin between platform i and platform j at time t, $p_i(t)$ and $p_j(t)$ represent the position vectors of the two platforms, Δ_d represents the minimum safety distance threshold, and $\chi_{ij}(t)$ represents the compensation term caused by wind field and communication delay. When $C_{ij}(t)$ is close to zero, the system triggers a local re-planning or speed adjustment strategy to maintain the dual constraints of edge accuracy and safe separation.

Finally, in order to evaluate the track quality under multi-objective conditions, this paper constructs the comprehensive fitness function as follows.

$$F=\beta_1\frac{L_{path}}{L_{ref}}+\beta_2\frac{T_{mission}}{T_{ref}}+\beta_3\frac{\sum\Psi}{\Psi_{ref}}+\beta_4\frac{E}{E_{ref}}+\beta_5\frac{\Gamma}{\Gamma_{ref}} \tag{15}$$

where F represents the track fitness, L_{path} represents the total distance, $T_{mission}$ represents the total task time, $\sum \Psi$ represents the cumulative risk along the path, E represents the total energy consumption, Γ represents the deviation between the edge error and the overlap rate, L_{ref}, T_{ref}, Ψ_{ref}, E_{ref}, Γ_{ref} represent the corresponding reference values, and β_1 to β_5 are the weights. This formula emphasizes not the single shortest path, but the overall optimum between the normative coverage requirements, the safety obstacle avoidance requirements and the high frequency inspection efficiency.

Based on the above mechanism, the backtracking search is no longer only responsible for general parameter optimization tasks, but directly acts on unit relay, local re-planning and safe obstacle avoidance control in long-distance embankment inspection. The generated track can not only meet the continuous coverage of multi-aircraft collaboration, but also maintain high execution stability in the face of local obstacles, wind disturbances or link fluctuations, and provide quantifiable path quality indicators for subsequent result analysis.

4 Results

4.1 Confusion matrix

In order to test the classification and discrimination ability of the dike disease and risk point recognition model in the real inspection scene, 112 hidden danger samples and the corresponding normal samples that were manually reviewed and confirmed after the inspection of the 72 km embankment section were selected to construct a test set, and the statistics were carried out according to five categories: cracks, leakage, gully collapse, illegal encroachment and normal background. The results show that the model identifies 880 normal samples, 108 diseases and risk points correctly, and 4 misjudgments, including 2 missed detection of fine cracks, 1 misclassification of weak thermal anomalies, and 1 false alarm of background texture. The overall discrimination results are stable. The proportion of the main diagonal in the confusion matrix is significantly higher than that of the off-diagonal area, indicating that the improved detection model has a strong ability to distinguish different types of embankment anomalies. In particular, the model still maintains high recognition consistency for cracks and leakage targets that are easy to interfere with each other. In order to visually present the correspondence between the predicted results of each category and the true labels, Fig. 5 shows the confusion matrix heatmap of the long-distance embankment inspection samples.

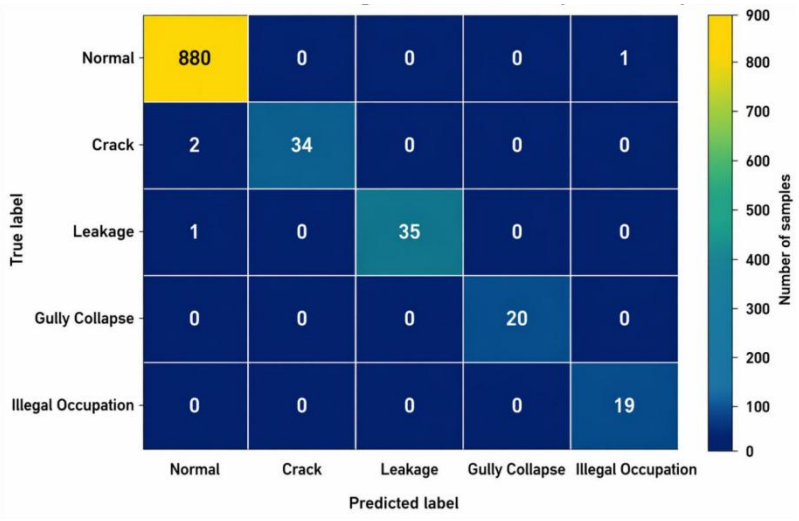

Figure 5: Heat map of confusion matrix for samples of long distance embankment inspections

As can be seen from Figure 5, the main diagonal area is continuously highlighted, which indicates that the model has a high classification hit rate on the five types of samples: normal background, crack, leakage, gully collapse and illegal encroachment. There is almost no cross misclassification between illegal encroachment and gully collapse, which indicates that the extraction of structural boundary and geometric shape features is sufficient. A small amount of deviation is mainly concentrated in the early fine cracks and low amplitude thermal anomaly samples, reflecting that the weak disease still has local response attenuation under complex background, but it does not affect the stable identification under normal inspection conditions.

4.2 Analysis of cruise height variation and specification compliance during inspection

The height of the patrol inspection directly affects the ground resolution, the accuracy of the edge connection and the quality of the full section coverage, so the height records of the fixed-wing mainline operation and the multi-rotor fine operation are analyzed in a unified way. The results show that the average measured value of the fixed wing at 100 m design altitude is 99.2 m, and the standard deviation is 1.8 m. The average measured value of the multi-rotor at the design height of 15 m is 14.7 m, and the standard deviation is 0.4 m, which are controlled within the allowable deviation range of the code. The time-division observation shows that the fluctuation of navigation height mainly appears in the stage of transition connection, local lifting and fine mining of embankment foot, but there is no continuous exceeding of the limit, which indicates that the cluster control terminal has a strong stabilizing effect on height maintenance and local correction. In order to clearly show the correspondence between the variation of the aero-height and the normative threshold, Fig. 6 shows the strip-like distribution of the deviation of the aero-height between the fixed and multi-rotor wings.

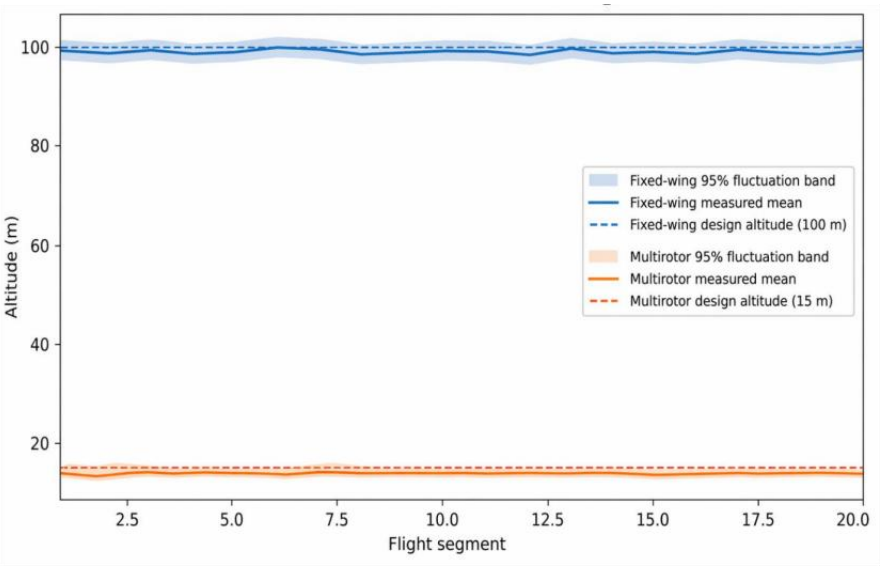

Figure 6: Strip distribution of fixed and multirotor airheight deviation

In Fig. 6, the middle axis represents the measured mean, the upper and lower boundaries represent the 95% fluctuation interval, and the dashed line represents the normative limit. It can be seen that most of the samples are distributed inside the safety belt, and a small number of disturbance points return quickly after a short correction, which proves that a good

closed-loop response is maintained between the cooperative path planning and the flight control. The navigation height control stability supports the sampling target of GSD not higher than 12 mm for macro inspection and 3 mm for fine inspection, and reduces the geometric error caused by edge drift.

4.3 Accuracy

The accuracy reflects the overall discrimination accuracy of the model for all samples, which can directly reflect the stable output ability of the normalized inspection link in multi-category scenarios. Based on the complete test set of 72 km embankment, this paper makes a unified comparison of five detection methods, including the single visible light detection model, the visible-infrared dual-mode model, the three-mode model with the removal of the direction sensitive attention module, the three-mode model with the removal of the risk guidance classification module, and the improved detection model in this paper. The results show that with the gradual enhancement of the input mode and the discrimination mechanism, the overall discrimination ability of the model for cracks, seepage, gully collapse and illegal encroachment continues to improve. Especially under the conditions of complex slope texture, shadow occlusion and low amplitude thermal anomalies, the single-modal model is more likely to misjudge the background disturbance as an anomaly. Although the dual-modal model improves the detection of thermal anomalies, the control of spatial boundaries is still not stable enough. After adding direction sensitive features, the continuous recognition effect of strip cracks and gully edges is significantly improved. After the risk-guided classification is further introduced, the overall classification stability of the model for multi-class anomalies continues to improve. To show the differences in the overall discrimination correctness of different models, Table 3 presents the accuracy comparison results.

Table 3: Accuracy comparison results of different detection models

Method	Input Data Composition	Accuracy / %	Result Description
Single visible-light detection model	Visible-light images	96.84	Performs well in detecting obvious cracks, but remains less stable for weak thermal anomalies and complex backgrounds
Visible-light–infrared dual-modal model	Visible light + infrared	97.13	Thermal anomaly recognition is improved, but spatial boundary discrimination still shows fluctuations
Three-modal model without direction-sensitive attention	Visible light + infrared + elevation	97.46	Multi-source representation is enhanced, but the continuity of slender crack detection is insufficient
Three-modal model without risk-guided grading	Visible light + infrared + elevation	97.72	Overall classification is relatively stable, but the ranking ability for multiple risk categories is weaker
Improved detection model	Visible light + infrared + elevation + direction-sensitive attention + risk-guided grading	98.21	Delivers the strongest overall discrimination capability and maintains the highest stability under complex backgrounds

It can be seen from Table 3 that the improved detection model achieves the highest accuracy of 98.21%, indicating that it has stronger advantages in multi-source information collaborative expression and complex background classification stability. Compared with the

single visible light detection model, the accuracy is improved by 1.37 percentage points, indicating that the infrared temperature difference information and elevation disturbance characteristics have an obvious supporting role in complex scene recognition. Compared with the model that removes the risk guidance classification module, the accuracy continues to increase by 0.49 percentage points, indicating that the risk guidance mechanism not only affects the subsequent sorting process, but also enhances the stability of the classification level for the discrimination of boundary samples. Combined with the above confusion matrix results, it can be further judged that the accuracy improvement is not an accidental improvement of a certain class of samples, but a synchronous optimization of the overall recognition link on multi-class targets.

4.4 Precision

Precision represents the proportion of abnormal samples detected by the model that really belong to the target class, and this indicator is directly related to the level of false alarm control. For the regular inspection of embankments, a low accuracy rate will lead to the misjudgment of normal slope texture, plant shadow, water reflection and local wet areas as risk points, thereby increasing the burden of manual review, extending the review time of results, and weakening the credibility of the hidden danger account. Based on the same test set, this paper continues to compare the precision rate changes of the five detection methods. The results show that the single visible light detection model is prone to false positives in areas with complex normal background textures, and the accuracy is the lowest. With the addition of IR, the discrimination of samples related to covert leakage is enhanced, but local temperature difference perturbations may still be amplified. After the introduction of elevation information, the model is more stable in identifying the collapse of the trench and the geometric anomaly of the slope, and the false alarm rate decreases. After further adding the direction sensitive attention module, the crack texture extended along the embankment is more clearly expressed, and the ability to distinguish cracks from background cracks and weed texture is significantly improved. To show the difference of each model in false alarm suppression, Table 4 presents the accuracy comparison results.

Table 4: Comparison results of precision rates for different detection models

Method	Input Data Composition	Precision / %	False-Positive Control Performance
Single visible-light detection model	Visible-light images	95.46	Easily affected by shadows, vegetation textures, and water-surface reflections
Visible-light–infrared dual-modal model	Visible light + infrared	96.11	Thermal anomaly discrimination is enhanced, but local temperature fluctuations may still lead to false positives
Three-modal model without direction-sensitive attention	Visible light + infrared + elevation	96.38	Shows good control over geometric anomalies, but false positives still occur along slender crack boundaries
Three-modal model without risk-guided grading	Visible light + infrared + elevation	96.87	Detection of multiple anomaly types is relatively stable, but lacks subsequent risk-screening constraints
Improved detection model	Visible light + infrared + elevation + direction-sensitive attention + risk-guided grading	97.52	Achieves the lowest false-positive rate and is the most suitable for normalized inspection output

It can be seen from Table 4 that the improved detection model maintains the highest value in the accuracy index, indicating that after multi-source information fusion, the misjudgment of the model for grass texture, slope shadow and local wet zone is significantly reduced. Compared with the single visible light detection model, the accuracy is increased by 2.06 percentage points, indicating that the multi-modal features and the direction constraint mechanism have a significant effect on the control of false alarms. Compared with the three-modal model that removes the risk-guided classification module, the improved model still increases by 0.65 percentage points, indicating that the risk-guided classification not only takes part in the ranking of results, but also can further reduce the false positive samples at the output. For the normalization inspection, this improvement has direct engineering significance, because after the number of false positives decreases, the manual review path is shorter, the hidden danger account generation efficiency is higher, and the output is closer to the stable state required by the specification.

4.5 Recall

Recall represents the proportion of all real abnormal samples successfully detected by the model, which is the core index to measure the ability of missing detection control. For dike engineering, the recall rate has more engineering significance than the pure appearance classification effect, because once the fine cracks, early leakage and local gully are missed in the inspection, they will directly affect the integrity of the danger detection, and then affect the subsequent disposal time. Based on the complete test set of 72 km embankment sections, this paper compares the recall of the five detection methods. The results show that the single visible light detection model can maintain a good detection of obvious apparent diseases, but it is insufficient to cover the early hidden dangers with thermal anomalies and unobvious textures. After introducing infrared into the bimodal model, the detection rate of leakage samples is significantly improved, but there are still local missed detection for the collapse of gullies with fine boundaries. After adding the elevation feature, the three-mode model has a stronger response to the risk points related to slope relief and terrain disturbance. After further superimposing direction sensitive attention, the continuous detection ability of the model for slender cracks and strip scour continues to be enhanced. To show the difference in the true anomaly detection ability of different methods, Table 5 presents the recall comparison results.

Table 5: Recall comparison results for different detection models

Method	Input Data Composition	Recall / %	Missed-Detection Control Performance
Single visible-light detection model	Visible-light images	96.03	Effective for obvious surface defects, but early concealed anomalies are still more likely to be missed
Visible-light–infrared dual-modal model	Visible light + infrared	96.71	Detection of seepage-related targets is improved, but some locally bounded anomalies are still overlooked
Three-modal model without direction-sensitive attention	Visible light + infrared + elevation	97.02	Detection of terrain-related anomalies is enhanced, but the continuous response to slender targets remains insufficient
Three-modal model without risk-guided grading	Visible light + infrared + elevation	97.24	Overall recall is relatively high, but the ranking and screening ability for weak targets is limited
Improved detection model	Visible light + infrared + elevation + direction-sensitive attention	97.64	Achieves the best detection completeness and is the most suitable for high-frequency inspection tasks

	+ risk-guided grading		

As can be seen from Table 5, the recall rate of the improved detection model reaches 97.64%, indicating that it can more completely capture multiple types of levee diseases and risk points in the normal inspection scene. Compared with the single visible light detection model, the recall rate is increased by 1.61 percentage points, indicating that thermal infrared and elevation information have obvious compensation effects on early hidden anomaly detection. Compared with the model that removes the risk-guided grading module, the recall rate continues to increase by 0.40 percentage points, indicating that the risk-guided mechanism helps to enhance the retention probability of weak targets in the output stage. For the high-frequency inspection task of long-distance levees, the continuous improvement of the recall rate means that the system can find the risk points more comprehensively, which is more in line with the requirements of the integrity of hidden danger discovery under the guidance of norms, and also provides a more sufficient data basis for subsequent manual review, risk classification and disposal scheduling.

4.6 F1-score

F1 value is the harmonic mean of precision and recall, which can reflect the balance between false alarm suppression and miss detection control. For the embankment disease identification task, it is difficult to meet the needs of normalized inspection driven by specifications by only improving the precision rate at the expense of recall rate, or only increasing the recall rate and introducing a large number of false positives. In the experiment, the F1 value of the improved detection model reaches 97.58%, which is significantly higher than 95.74% of the single visible light detection model and 96.28% of the model without direction sensitive attention module. To show the differences in the comprehensive recognition stability of different models, Fig. 7 shows the F1-value radar comparison plot。

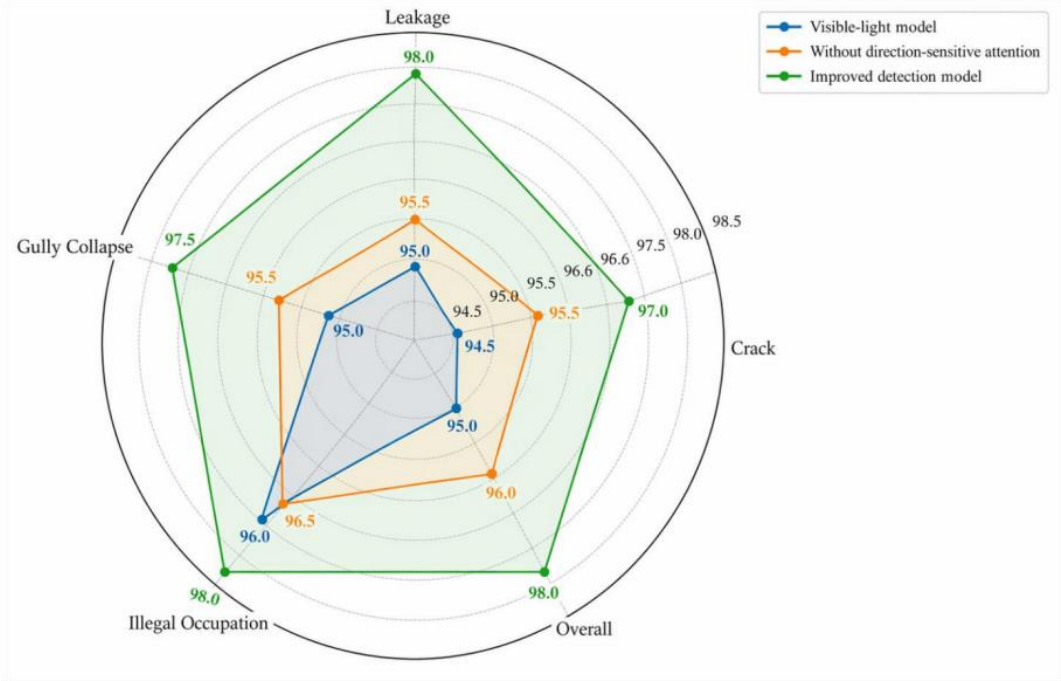

Figure 7: Radar comparison plot of F1 values for different detection models

Fig. 7 shows that the improved detection model is superior to the comparison model in terms of comprehensive balance, and its advantages are not concentrated on a single category,

but maintain a relatively stable F1 distribution in the four categories of targets: cracks, leakage, gully collapse and illegal encroachment. The results show that the adjustment of the model structure not only improves the single index, but also improves the overall consistency of the identified link. For the embankment inspection system that needs long-term operation and periodic reuse, this balance is more meaningful than a single high value.

4.7 Intersection and union ratio

The intersection over Union (MIU) is used to measure the spatial overlap degree between the predicted region and the real labeled region, which is an important index to evaluate the accuracy of disease localization. In the embankment inspection, the intersection over union ratio not only relates to whether the target frame accurately covers the cracks, gullies and leakage areas, but also directly affects the geometric reliability of the hidden danger size measurement, location labeling and subsequent results filing. On the basis of the above classification results maintaining high stability, the intersection over union ratio can continue to reflect whether the model is accurate in characterizing the abnormal boundary, so as to further test the reliability of the recognition results at the spatial location level. Experimental results show that the average intersection over union ratio of the improved detection model reaches 96.18%, which is higher than 93.87% of the single visible light detection model and 94.52% of the model removing cross-scale recombination module. To show the distribution of different target categories in spatial positioning accuracy, Fig. 8 shows the violin distribution of the intersection and union ratio of various diseases.

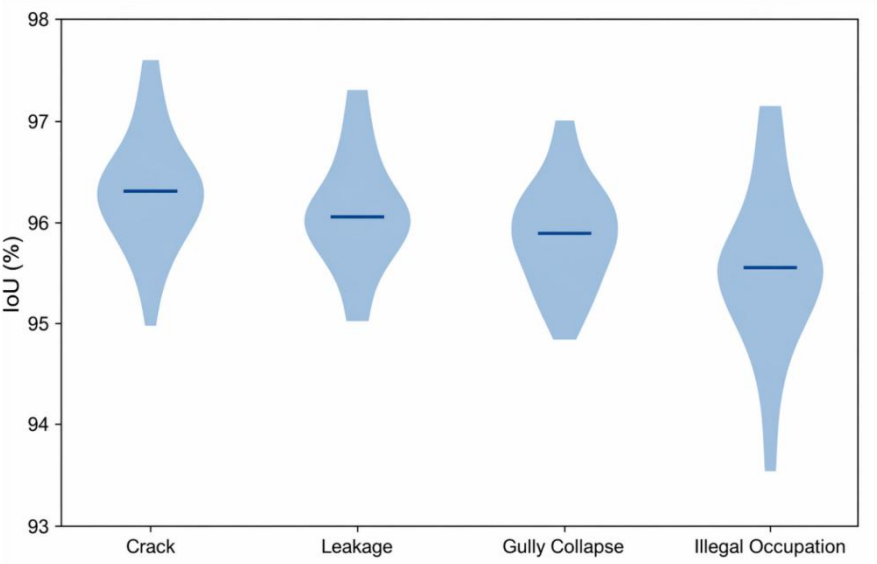

Figure 8: Violin plot of intersection and union ratio distribution of various types of dike diseases

As can be seen from Fig. 8, the distribution of the improved detection model on the three types of targets, crack, leakage and gully collapse, is relatively concentrated, and the median value remains in a high range, indicating that the positioning of the main disease boundaries by the model is relatively stable. The distribution of illegal encroachment is slightly wider, mainly related to the large difference in target morphology, but the overall is still at a high intersection and union ratio level. The results show that the model output can support the subsequent coordinate labeling, risk area delineation and specification production.

4.8 Five-fold cross-validation

In order to test the generalization ability and result stability of the improved detection model under different sample division conditions, the five-fold cross validation method was used to train and test the complete data set repeatedly. Each fold keeps the proportion of categories basically the same, and outputs five indicators: accuracy, precision, recall, F1 score and intersection/union ratio. The results show that the fluctuations among folds are small, which indicates that the model has strong adaptability on long-distance embankment inspection data. In order to present the stability results under the 50-fold verification condition, Table 6 shows the statistical values of the main performance indexes of each fold.

Table 6: Statistical table of five-fold cross-validation results

Fold	Accuracy / %	Precision / %	Recall / %	F1-score / %	IoU / %
Fold 1	97.94	97.31	97.42	97.36	95.92
Fold 2	98.08	97.48	97.57	97.52	96.11
Fold 3	98.16	97.55	97.62	97.58	96.18
Fold 4	97.98	97.39	97.46	97.42	95.97
Fold 5	98.11	97.52	97.64	97.58	96.18

As can be seen from Table 6, under the five-fold cross validation, all the indicators remain high, the accuracy rate is distributed between 97.94% and 98.16%, the fluctuation range of the precision and recall rate is not large, and the F1 value and the intersection and union ratio remain stable. The results show that the model does not rely too much on a fixed training set division, but can maintain good recognition and positioning ability under different data subsets, which is suitable for normal and multi-cycle embankment inspection tasks.

5 Discussion

The normalized inspection system of UAV swarm constructed in this paper shows strong specification adaptability and calculation cooperativity in the long-distance embankment engineering scenario. Compared with traditional manual inspection and single UAV segmented operation, the advantages of the system are not only reflected in the results of completing 100-kilometer full-section inspection in 5.2 hours, improving efficiency by 32 times and reducing human investment by 80%, but also reflected in the unification of data link, decision link and achievement link. Multi-source sensor input enables visible light texture, thermal anomaly information and elevation disturbance to participate in the discrimination under the same coordinate reference. The improved detection model incorporates cracks, leakage, gully and illegal encroachment into the same recognition framework. Therefore, the minimum crack identification width reaches 1.8mm, and the hidden danger account, DOM, DSM and inspection report can be generated synchronously, indicating that the system has been transformed from a single-point identification tool to a full-process computing platform for specification execution. At the same time, the results also show that the system can still maintain stable output under complex terrain, dense structure area and high frequency re-mining conditions, which is particularly critical for monthly inspection in non-flood season and weekly inspection in flood season. When discussing the significance of this research, we should not only understand it as equipment replacing human labor, but also see the coupling relationship between cluster scheduling, model identification, quality verification and result archiving. It is this scene-oriented collaborative computing mechanism that makes long-distance embankment inspection have the engineering value of

standardization, continuity and traceability. This also shows that the new code has a stable foundation for popularization in similar river basin projects.

6 Conclusions

Around the requirements of long-distance embankment regular inspection under the constraints of SL/T860-2026, this paper builds a technical system that links cluster platform configuration, three-level collaborative route planning, multi-source data processing, improved detection model, backtracking search scheduling and results archiving, and completes the whole process verification in a 72km secondary embankment scenario. The results show that the system can achieve 100% coverage, 100% data archiving and 100% code compliance verification. A single full section inspection takes only 5.2 hours, and it can identify the cracks and hidden leakage abnormalities of the embankment with a width of not less than 1.8mm, which indicates that the UAV cluster has a strong engineering deployment ability in the specification-driven scenario. At the same time, this paper still has some limitations. The sample sources mainly focus on a single basin embankment, and the long-term continuous verification under complex meteorological conditions is still insufficient. The deep semantic alignment between thermal infrared, point cloud and visible light is still dominated by task-level fusion, and a stronger end-to-end joint learning mechanism has not been formed. There is still room for expansion of the linkage depth between spatial collaboration, edge computing power and digital twin platform. The subsequent research will build a larger training set for cross-basin multi-dike samples to enhance the generalization ability of the model under different hydrogeomorphological conditions, and introduce lightweight deployment, online incremental learning and time series risk prediction methods to further improve the real-time performance, continuity and intelligent decision-making ability of the regular inspection. At the same time, it is necessary to improve the cross-regional airspace scheduling interface and multi-machine adaptive fault tolerance mechanism, so that the inspection results can more efficiently serve the integration application of embankment safety management platform and early warning link.

About the Author

Jiang Lijun graduated from Zhejiang University of Water Resources and Electric Power in 2006, majoring in Water Conservancy and Hydropower Engineering. She is currently Director of the Pujiang County Water Conservancy Planning, Construction and Quality & Safety Management Center. She has formulated a series of measures and opinions, which have effectively standardized the construction and management of small-scale water conservancy projects in Pujiang and promoted the high-quality development of local water conservancy projects.

References

[1] Puente-Castro A, Rivero D, Pazos A, et al. A review of artificial intelligence applied to path planning in UAV swarms[J]. Neural Computing and Applications, 2022, 34(1): 153-170.

[2] Saeed R A, Omri M, Abdel-Khalek S, et al. Optimal path planning for drones based on swarm intelligence algorithm[J]. Neural Computing and Applications, 2022, 34(12):

10133-10155.

[3] Jeong E, Seo J, Wacker J P. UAV-aided bridge inspection protocol through machine learning with improved visibility images[J]. Expert Systems with Applications, 2022, 197: 116791.

[4] Luna M A, Ale Isaac M S, Ragab A R, et al. Fast multi-UAV path planning for optimal area coverage in aerial sensing applications[J]. Sensors, 2022, 22(6): 2297.

[5] Espinosa Peralta P, Luna M A, de la Puente P, et al. Performance analysis of localization algorithms for inspections in 2D and 3D unstructured environments using 3D laser sensors and UAVs[J]. Sensors, 2022, 22(14): 5122.

[6] Flores Peña P, Luna M A, Ale Isaac M S, et al. A proposed system for multi-UAVs in remote sensing operations[J]. Sensors, 2022, 22(23): 9180.

[7] Shafiq M, Ali Z A, Israr A, et al. Convergence analysis of path planning of multi-UAVs using max-min ant colony optimization approach[J]. Sensors, 2022, 22(14): 5395.

[8] Avdelidis N P, Tsourdos A, Lafiosca P, et al. Defects recognition algorithm development from visual UAV inspections[J]. Sensors, 2022, 22(13): 4682.

[9] Gugan G, Haque A. Path planning for autonomous drones: Challenges and future directions[J]. Drones, 2023, 7(3): 169.

[10] Cetinsaya B, Reiners D, Cruz-Neira C. From pid to swarms: A decade of advancements in drone control and path planning-a systematic review (2013–2023)[J]. Swarm and Evolutionary Computation, 2024, 89: 101626.

[11] Katkuri A V R, Madan H, Khatri N, et al. Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review[J]. Array, 2024, 23: 100361.

[12] Ejaz N, Choudhury S. Computer vision in drone imagery for infrastructure management[J]. Automation in Construction, 2024, 163: 105418.

[13] Luna M A, Molina M, Da-Silva-Gomez R, et al. A multi-UAV system for coverage path planning applications with in-flight re-planning capabilities[J]. Journal of Field Robotics, 2024, 41(5): 1480-1497.

[14] Boukabou I, Kaabouch N. Electric and magnetic fields analysis of the safety distance for UAV inspection around extra-high voltage transmission lines[J]. Drones, 2024, 8(2): 47.

[15] Hansen J G, de Figueiredo R P. Active object detection and tracking using gimbal mechanisms for autonomous drone applications[J]. Drones, 2024, 8(2): 55.

[16] Santos T, Cunha T, Dias A, et al. UAV visual and thermographic power line detection using deep learning[J]. Sensors, 2024, 24(17): 5678.

[17] Takaya K, Ohta H, Shibayama K, et al. Construction of power line inspection system

using a quadrotor helicopter[J]. IEEJ Transactions on Electrical and Electronic Engineering, 2024, 19(11): 1803-1816.

[18] Quadt T, Lindelauf R, Voskuijl M, et al. Dealing with multiple optimization objectives for UAV path planning in hostile environments: A literature review[J]. Drones, 2024, 8(12): 769.

[19] Panowicz R, Stecz W. Robust Optimization Models for Planning Drone Swarm Missions[J]. Drones, 2024, 8(10): 572.

[20] Adoni W Y H, Fareedh J S, Lorenz S, et al. Intelligent swarm: Concept, design and validation of self-organized UAVs based on leader–followers paradigm for autonomous mission planning[J]. Drones, 2024, 8(10): 575.