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

Research on intelligent diagnosis and risk early warning model of multimodal pathological data of gastrointestinal mucosal lesions

Wenjing Fu[1], *Dake Yang*[3], *Wei Wang*[4] *and Jing Xia*[1,2,*]

[1] School of Medicine, Hainan Vocational University of Science and Technology, Haikou 571126, Hainan, China
[2] School of Pharmacy, China Medical University, Shenyang 110122, Liaoning, China
[3] Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, Shanghai, China
[4] School of Basic Medicine, Xinjiang Medical University, Urumqi 830017, Xinjiang, China

SUMMARY: *The fragility of post-earthquake public health environment has significantly increased the risk of gastrointestinal mucosal injury and infection spread. It is urgent to establish intelligent analysis methods that take into account recognition accuracy, early warning timeliness and complex data adaptability. In this paper, a multi-modal intelligent diagnosis and risk warning framework fusing pathological images, clinical texts and detection indicators is constructed for gastrointestinal mucosal lesions. Through standardized preprocessing, convolutional visual coding, medical text semantic representation, structured indicator embedding, shared space alignment, modal attention fusion and dual-branch collaborative decision making of lesion recognition-risk discrimination. The linkage output of lesion classification, risk score and hierarchical early warning was realized. The experimental results based on 1620 case samples show that the Accuracy, F1-score and AUC of the proposed method reach 93.2%, 92.5% and 0.964, respectively. The risk warning accuracy reaches 90.8%, the recall rate of Level IV cases reaches 94.1%, and the average warning advance time reaches 23.7 hours. The results of this study provide reliable technical support and practical significance for early screening, early diagnosis and stratified intervention of gastrointestinal mucosal lesions after earthquake.*

KEYWORDS: *Gastrointestinal mucosal lesions; Multi-modal fusion; Intelligent diagnosis; Risk warning*

1 Introduction

After an earthquake, the affected people are often faced with multiple impacts of environmental mutation, sudden increase of survival pressure and continuous uncertainty [1]. The destruction of living space, the temporary interruption of social support network, and the continuous stimulation of individual cognitive system by casualty information can easily induce acute stress disorder, and gradually evolve into chronic anxiety in a long period of high pressure. Once the anxiety disorder is not recognized in time, it will not only weaken the post-disaster population's behavior regulation ability and willingness to help, but also affect sleep, diet, immune function and the level of chronic disease control, and eventually cause the double deterioration of psychological and physical health [2, 3]. Because of this, accurate

*abcjx133@126.com
https://doi.org/10.65102/is2026740
Published: 30 April, 2026

identification modeling of anxiety disorders has become an important part of post-disaster medical monitoring, which helps to complete the screening of high-risk groups before the full manifestation of symptoms, and to buy valuable time for stratified intervention and psychological support [4].

Accompanied by psychological risks, the post-earthquake public health environment is rapidly weakened. Damage to infrastructure will lead to interruption of water supply system, out-of-control food storage and transportation conditions, and decline of sewage disposal capacity. Pathogenic microorganisms are more likely to spread into the human digestive tract through contaminated water, food and contact, and directly invade the gastrointestinal mucosa, causing inflammation, erosion, ulcers and even more serious infectious lesions [5, 6]. As an important interface between body barrier and immune response, gastrointestinal mucosa plays a prominent sentinel role in the spread chain of post-disaster infectious diseases [7]. If the degree and evolution trend of mucosal injury cannot be identified as early as possible, local lesions are likely to further induce the spread of mass infection and the crowding out of medical resources. It can be seen that the construction of an accurate identification and risk early warning model for gastrointestinal mucosal lesions is a key support for early diagnosis and treatment, blocking the deterioration of the epidemic, and improving the efficiency of post-disaster health emergency response [8, 9].

At present, the diagnosis of gastrointestinal mucosal lesions still mostly relies on single pathological image observation or empirical judgment. In the face of realistic situations such as complex case sources, overlapping lesion types and incomplete clinical information after earthquake, traditional methods have shortcomings in accuracy, timeliness and stability [10]. With the development of digital pathology, natural language processing, multimodal fusion learning and risk prediction technology, collaborative analysis of heterogeneous information such as pathological images, clinical texts and laboratory indicators has provided a new technical path for intelligent diagnosis and early warning of gastrointestinal mucosal lesions [11-13]. Based on this, this paper focuses on the multimodal pathological data of gastrointestinal mucosal lesions, and constructs an intelligent analysis framework for lesion recognition and risk early warning, in order to provide reliable support for the screening of high-risk populations and rapid clinical decision-making after disasters.

2 Theoretical basis and related research

2.1 Medical basis and computer representation mechanism of multimodal pathological analysis of gastrointestinal mucosal lesions

Gastrointestinal mucosa is an important interface between digestion and absorption and barrier defense, and its structural integrity is directly related to local inflammation control, immune response balance and pathogen blocking ability. In the post-earthquake environment, contaminated water, spoiled food and decreased sanitary conditions can significantly increase the risk of gastrointestinal infection. Bacteria, toxins and inflammatory mediators act together on the mucosal epithelium, which can induce continuous pathological changes such as congestion, edema, erosion, ulcer and glandular structure disorder [14]. Some patients only show mild mucosal injury in the early stage, but if they are not recognized in time, the lesions may continue to develop along the path of inflammatory diffusion, barrier damage and functional imbalance, and be coupled with clinical manifestations such as fever, diarrhea, bleeding and nutrient absorption disorders [15, 16]. Therefore, the evaluation of gastrointestinal mucosal lesions should not be limited to a single morphological observation, but should pay attention to histological characteristics, clinical symptoms and laboratory

index fluctuations.

From the perspective of computer representation, the core of multimodal pathology analysis is to transform information from different sources and scales into a unified expression that can be calculated, fused and discriminable. Pathological images can provide visual features such as glandular arrangement, cell density, nuclear morphology, inflammatory infiltration and mucosal boundary. Clinical text records carry symptom descriptions, medical history clues, diagnostic opinions and processing processes. Examination indicators can further reflect the intensity of infection, the level of inflammation and the metabolic state of the body [17]. Through image feature extraction, text semantic coding and structured index embedding, a cross-modal feature space covering morphological information, semantic information and numerical information can be established. With the help of feature alignment, correlation learning and risk mapping mechanism, the internal law of the evolution of gastrointestinal mucosal lesions from local damage to overall risk can be depicted [18-20]. This representation method not only helps to improve the accuracy of lesion recognition, but also provides a reliable data basis for subsequent risk early warning scoring and high-risk case screening.

2.2 Research status of intelligent diagnosis and risk early warning of gastrointestinal mucosal lesions

In recent years, the intelligent diagnosis of gastrointestinal mucosal lesions has gradually shifted from traditional manual movie-reading assistance to data-driven automatic recognition and analysis. Early related studies mostly focused on the independent interpretation of endoscopic images, pathological sections or single laboratory indicators, and used image segmentation, texture analysis and shallow machine learning methods to classify and identify lesions such as inflammation, erosion, ulcer and dysplasia [21]. Such methods have achieved certain results on standardized datasets, but limited by feature expression ability and data heterogeneity, the diagnostic stability is still insufficient in the face of complex lesion morphology, fuzzy boundary area and coexistence of multiple types of lesions [22].

With the development of deep learning technology, convolutional neural networks, visual converters and transfer learning methods have been widely used in gastrointestinal pathological image analysis, showing stronger feature extraction capabilities in gland structure recognition, cell morphology discrimination and lesion area localization [23]. At the same time, natural language processing technology has been introduced into the semantic mining of clinical medical records, pathological reports and follow-up records to extract key information such as symptom development, diagnosis and treatment process and risk description. Some studies have tried to jointly model image features and text features, and carry out multimodal analysis by combining inflammatory indicators, infection indicators and biochemical indicators, so that intelligent diagnosis is gradually extended from single visual recognition to cross-modal comprehensive discrimination [24, 25]. This change indicates that the study of gastrointestinal mucosal lesions is evolving from "disease recognition by looking at pictures" to "multi-source information collaborative decision-making".

Compared with intelligent diagnosis, risk early warning research started slightly later, and the existing work mainly focuses on disease progression prediction, postoperative complication assessment and high-risk screening of digestive tract tumors. In terms of specific methods, logistic regression, random forest, gradient boosting model or temporal neural network are often used to comprehensively analyze the changes in clinical indicators, lesion grading results and past medical history of patients to generate risk scores or early warning levels [26]. These studies provide a technical basis for early screening, but there are

still several deficiencies. On the one hand, many early warning models rely on structured data, and do not make full use of pathological image details and text semantic information. On the other hand, the model is mostly built for conventional medical scenarios, and does not consider enough the actual needs of complex case sources, rapid infection transmission and missing data in post-earthquake environment, which limits the generalization ability and emergency applicability.

In general, the existing research has laid a foundation for the intelligent recognition and risk assessment of gastrointestinal mucosal lesions, but there is still a large room for improvement in multimodal deep fusion, linkage modeling of diagnostic results and risk evolution, and adaptation of high-risk scenarios after disasters, which also constitutes the practical basis for further research in this paper.

3 Construction of intelligent diagnosis and risk early warning model for multimodal pathological data of gastrointestinal mucosal lesions

3.1 Multi-modal pathological data acquisition integration and standardized preprocessing method

Intelligent diagnosis and risk warning of gastrointestinal mucosal lesions rely on high-quality, multi-source and collaboratively analyzed data. In the post-earthquake medical environment, the sources of cases are complex, the collection conditions fluctuate obviously, data missing and noise interference are more likely to occur, and the subsequent feature coding and risk discrimination results are easily affected. Therefore, this study constructs a multi-modal data acquisition and integration process around three types of core information: pathological images, clinical texts and detection indicators, and designs a system in sample-level alignment, quality control, standardization processing and anomaly removal to ensure the consistency of input data in structure, semantic and numerical scales. The multimodal pathology data acquisition and preprocessing process is shown in Figure 1.

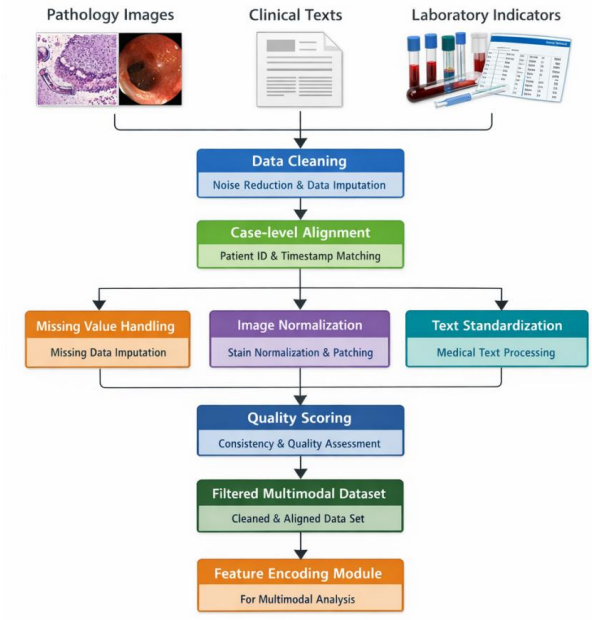

Figure 1: Flow chart of multimodal pathology data acquisition and preprocessing

In the data acquisition stage, the gastrointestinal mucosal lesion samples are represented as a ternary input set composed of image modality, text modality and indicator modality. For the ith case sample, it can be written as follows:

$$X_i = \{I_i, T_i, L_i\}, i = 1,2, \ldots, N \tag{1}$$

Among them, I^i represents the collection of digital pathological slices or pathological image blocks, T_i represents the medical record, pathological report and symptom description text corresponding to the case, and L_i represents the routine blood test, inflammatory indicators, biochemical indicators and infection-related test values. By unifying the case number, the visit timestamp and the specimen number, the corresponding association of different modal data at the case level was realized, and the cross-modal mismatch problem was reduced from the source. Considering that the data acquisition time and equipment conditions may be unstable in post-earthquake scenarios, an integrity marker matrix was established for all cases in the study to record the completeness of each modality, so as to provide a basis for subsequent missing completion and sample screening.

In terms of image preprocessing, the original pathological images usually have problems such as resolution differences, different shades of staining, more background redundancy, and local artifact interference. In order to improve the consistency of image input, we successively perform denoising, background cropping, color normalization and sliding window segmentation. The image block after cutting is denoted as pij, where j is the JTH image block corresponding to the ith case. To ensure the comparability of different samples in pixel distribution, the image intensity is normalized:

$$\hat{p}_{ij} = \frac{p_{ij} - \min(p_{ij})}{\max(p_{ij}) - \min(p_{ij}) + \varepsilon} \tag{2}$$

Here, $\hat{p}_{ij}$ is the normalized image block and ε is a tiny constant to prevent the denominator from being zero. After this step, the image data tends to be unified in brightness scale and local contrast, which is conducive to the stable extraction of key pathological features such as gland arrangement, cell density, inflammatory infiltration and mucosal boundary by subsequent models.

In terms of clinical text processing, raw medical records and pathological reports contain a large number of unstructured descriptions, which have problems such as differences in terminology expression, mixed acronyms and redundant statements. In order to improve the semantic quality of the text, this paper firstly cleans the text, eliminates irrelevant symbols, repeated fields and template-oriented expressions with no actual diagnostic significance, and then combines synonyms and near-sense descriptions into a unified label space through the standard mapping of medical terms. Then, text segments are segmented according to the dimensions of symptom information, past history, endoscopic observation, pathological opinion and treatment suggestion, and negative words, adverbs of degree and timeline information are retained to reduce semantic deviation. For each text record, both the original text and the structured label results are retained in the study, so that the subsequent feature coding can take advantage of both deep semantics and clinical rule information.

In terms of detection index processing, laboratory data are characterized by many dimensions, large dimension differences, and scattered missing points. If it is directly input into the model, the difference of numerical magnitude will cause the imbalance of feature contribution. In this regard, this paper uses the standard score transformation to standardize the JTH index:

$$z_{ij}=\frac{x_{ij}-\mu_j}{\sigma_j+\varepsilon} \tag{3}$$

Here, x_{ij} represents the original value of the ith case on the JTH detected index, μ_j and σ_j represent the mean and standard deviation of this index in the training set, respectively, and z_{ij} is the normalized result. For a small number of missing items, the imputation method based on the neighborhood of similar cases was used to complete the missing items. For the samples with too high missing proportion or missing key indicators, they are directly eliminated to avoid the error being amplified in the subsequent fusion stage.

The key to multimodal integration is case-level registration and sample quality assessment. In order to measure the overall usability of a single case sample, this paper constructs a quality scoring function:

$$Q_i=\alpha_1 q_i^{img}+\alpha_2 q_i^{txt}+\alpha_3 q_i^{lab} \tag{4}$$

Here,q_i^{img}, q_i^{txt} and q_i^{lab} represent the image clarity and coverage score, the text integrity and standardization score, and the detection index completeness and credibility score, respectively. α_1, α_2, α_3 are the corresponding weights, and meet the following requirements:

$$\alpha_1+\alpha_2+\alpha_3=1 \tag{5}$$

When Q_i is lower than the preset threshold, the sample will be marked as a low-quality sample and enter the manual review or rejection process. This mechanism can screen out obvious abnormal inputs before multimodal fusion, and enhance the reliability of the whole process data basis.

After the standardization of images, texts and indicators is completed, a unified data organization unit is established according to the case dimension, and an input format suitable for subsequent cross-modal feature coding is generated. On the one hand, this method improves the correspondence accuracy between different modalities, on the other hand, reduces the interference of noise, missing and heterogeneous differences in the complex scene after the earthquake on the model training, and lays a stable data foundation for the subsequent lesion identification and risk early warning analysis.

3.2 Cross-modal feature coding method for clinical text and detection indicators of pathological images

The formation and evolution of gastrointestinal mucosal lesions are simultaneously reflected in three levels: tissue morphology, clinical semantics and detection values. A single modality can only reveal local information, which is difficult to completely reflect the disease status and risk changes. In order to improve the representation ability of subsequent diagnosis and early warning analysis, this paper constructs a cross-modal feature coding method for pathological images, clinical texts and detection indicators. On the basis of maintaining the original information characteristics of each modality, the heterogeneous data are mapped into a unified shared feature space, and the semantic consistency between different modalities is enhanced by an alignment mechanism. The cross-modal feature coding and alignment structure are shown in Figure 2, and the description of different modal feature coding methods is shown in Table 1.

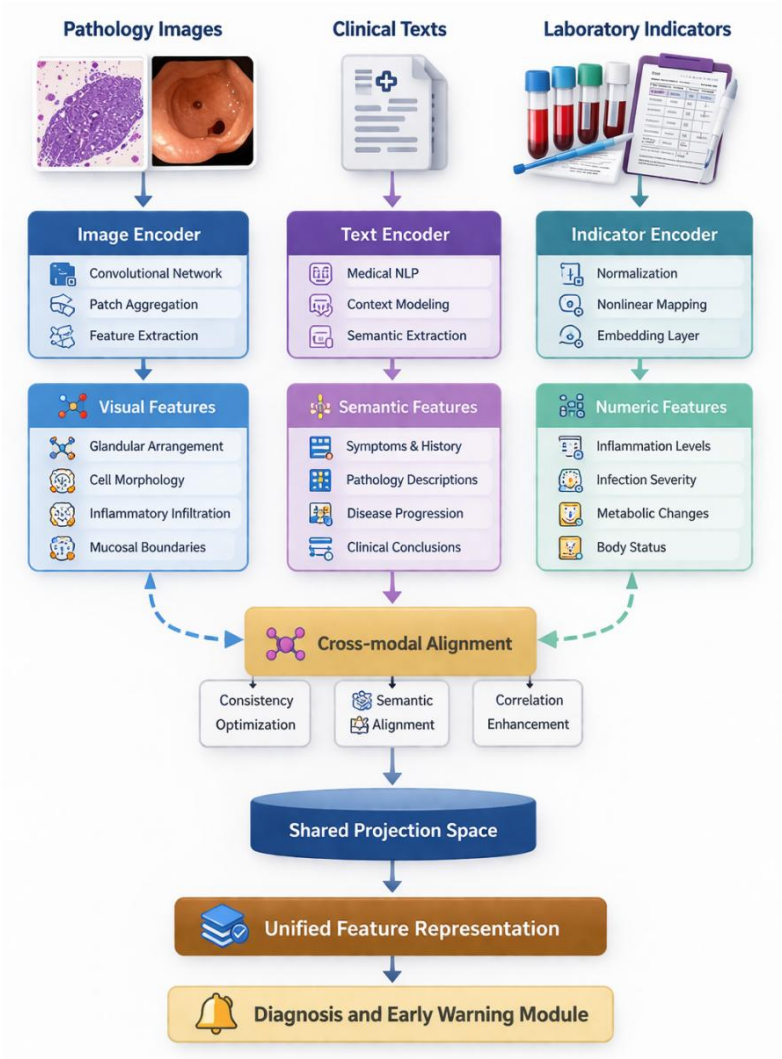

Figure 2: Structure diagram of cross-modal feature encoding and alignment

Table 1: Illustrative table of different modal feature encoding methods

Modality Type	Input Form	Encoding Method	Output Representation	Representation Focus
Pathological Image Modality	Digital pathological slides, image patches	Image encoder extracts local texture and spatial structural features	Visual Features	Gland arrangement, cellular morphology, inflammatory infiltration, mucosal boundaries
Clinical Text Modality	Medical records, pathological reports, symptom descriptions	Text encoder extracts contextual semantics and diagnostic information	Semantic Features	Symptom progression, medical history clues, pathological opinions, risk alerts
Laboratory Indicator Modality	Inflammatory indicators, biochemical indicators, infection-related numerical values	Linear mapping and nonlinear embedding encoding	Numeric Features	Inflammation level, infection intensity, metabolic changes, physiological status
Cross-modal Shared Representation	Encoding results from the three modalities	Shared-space projection and alignment learning	Unified Feature Representation	Modality consistency, complementary information integration, case-level joint representation

For pathological image modalities, this paper takes the standardized pre-processed image patches as input, and uses a convolutional feature extraction network to obtain visual representations such as local texture, gland arrangement, cell density and inflammatory infiltration. Let the image input of the ith case be I_i, then the image coding result can be expressed as follows:

$$v_i^{img} = E_{img}(I_i) \tag{6}$$

Here, $E_{img}(\cdot)$ represents the image encoder, and v_i^{img} is the high-dimensional visual feature vector of the image modality. Considering the uneven distribution of lesion regions in pathological images, we introduce a regional aggregation operation after encoding to summarize the local representations of multiple image blocks, so as to obtain a case-level image representation and reduce the disturbance of local noise on the global representation.

For the clinical text modality, medical record descriptions, pathology reports, and symptom records contain a large number of semantic clues related to the extent of lesions, infection status, and evolution trends. In this paper, the standardized text sequence is input into the medical text coding network, and the context semantic relationship and key information dependence are extracted to obtain the text semantic representation:

$$v_i^{txt} = E_{txt}(T_i) \tag{7}$$

Here, T_i represents the text sequence corresponding to the ith case, $E_{txt}(\cdot)$ is the text encoder, and v_i^{txt} is the text modal feature vector. The representation can cover the duration of symptoms, inflammation description, bleeding, infection hints, diagnosis and treatment opinions, and provide semantic support for subsequent feature alignment.

For the detection indicator modality, structured numerical data can supplement the information that cannot be directly reflected by pathological images and clinical texts from the perspectives of inflammation intensity, infection level, electrolyte status and metabolic changes. Let the normalized index vector of the ith case be Li, then its coding result can be written as follows:

$$v_i^{lab} = E_{lab}(L_i) = \phi(W_{lab}L_i + b_{lab}) \tag{8}$$

Here, W_{lab} and b_{lab} represent the linear mapping parameter and bias term, $\phi(\cdot)$ represents the nonlinear activation function, and vi_{lab} is the embedding representation of the detection index mode. After this transformation, laboratory data with large differences in original dimensions are mapped to a low-dimensional continuous space that is more suitable for joint analysis.

Due to the obvious differences between the three modalities in data form, scale structure and semantic level, it is easy to cause information conflict and contribution imbalance between modalities if the features are directly concatenation. To this end, in this paper, each modal feature is separately projected into a shared space of uniform dimension. For any modality $m \in \{img, txt, lab\}$, its shared representation is defined as follows:

$$h_i^m = W_m v_i^m + b_m \tag{9}$$

Here, W_m and b_m are the projection parameters of the corresponding modality, and h_i^m is the feature representation after mapping to the shared space. By unifying the dimension constraints, different modalities can establish correlations in the same representation space,

which lays the foundation for subsequent collaborative discrimination.

Based on this, this paper further designs a cross-modal alignment mechanism to enhance the consistent expression between different modalities of the same case. Considering the association between three types of features: image, text and metrics, the alignment loss is defined as follows:

$$L_{align} = \|h_i^{img} - h_i^{txt}\|_2^2 + \|h_i^{img} - h_i^{lab}\|_2^2 + \|h_i^{txt} - h_i^{lab}\|_2^2 \tag{10}$$

Here, $\|\cdot\|2$ denotes the two-norm. By compressing the distance between the cross-modal representations of the same case, the pathological morphology presented by the image, the clinical semantics described by the text and the physiological changes reflected by the indicators form a more stable joint expression in the shared space. In order to avoid excessive influence of a certain modality on the overall representation due to data quality fluctuations, this paper introduces a modal weight allocation strategy in the shared space to generate the final encoding result:

$$H_i = \beta_1 h_i^{img} + \beta_2 h_i^{txt} + \beta_3 h_i^{lab} \tag{11}$$

Here, H_i represents the cross-modal joint coding result of the ith case, β_1, β_2, β_3 are the modal weights and are satisfied:

$$\beta_1+\beta_2+\beta_3=1 \tag{12}$$

The joint encoding vector further improves the stability and transferability of the representation while retaining the complementary information of different modalities. After the above encoding and alignment processing, the visual, semantic and numerical information related to gastrointestinal mucosal lesions are integrated into a unified case-level feature representation, which provides high-quality input for the multimodal fusion intelligent diagnosis in the next section.

3.3 Multimodal fusion intelligent diagnosis method for lesion recognition and risk discrimination

After the cross-modal coding of pathological images, clinical texts and detection indicators is completed, how to transform the three kinds of heterogeneous information into a unified decision-making basis for lesion identification and risk discrimination becomes the key of intelligent diagnosis. Gastrointestinal mucosal lesions have the characteristics of complex morphological changes, obvious differences in clinical manifestations, and phased index fluctuations. Relying solely on a certain modality is easy to cause local judgment bias. Based on this, this paper constructs a multi-modal fusion intelligent diagnosis method for lesion recognition and risk discrimination. Through modal attention allocation, joint representation learning and dual-branch collaborative decision-making, the image morphological information, text semantic information and test numerical information form complementary support in the same framework. Figure 3 shows the collaborative decision-making mechanism of lesion identification and risk discrimination.

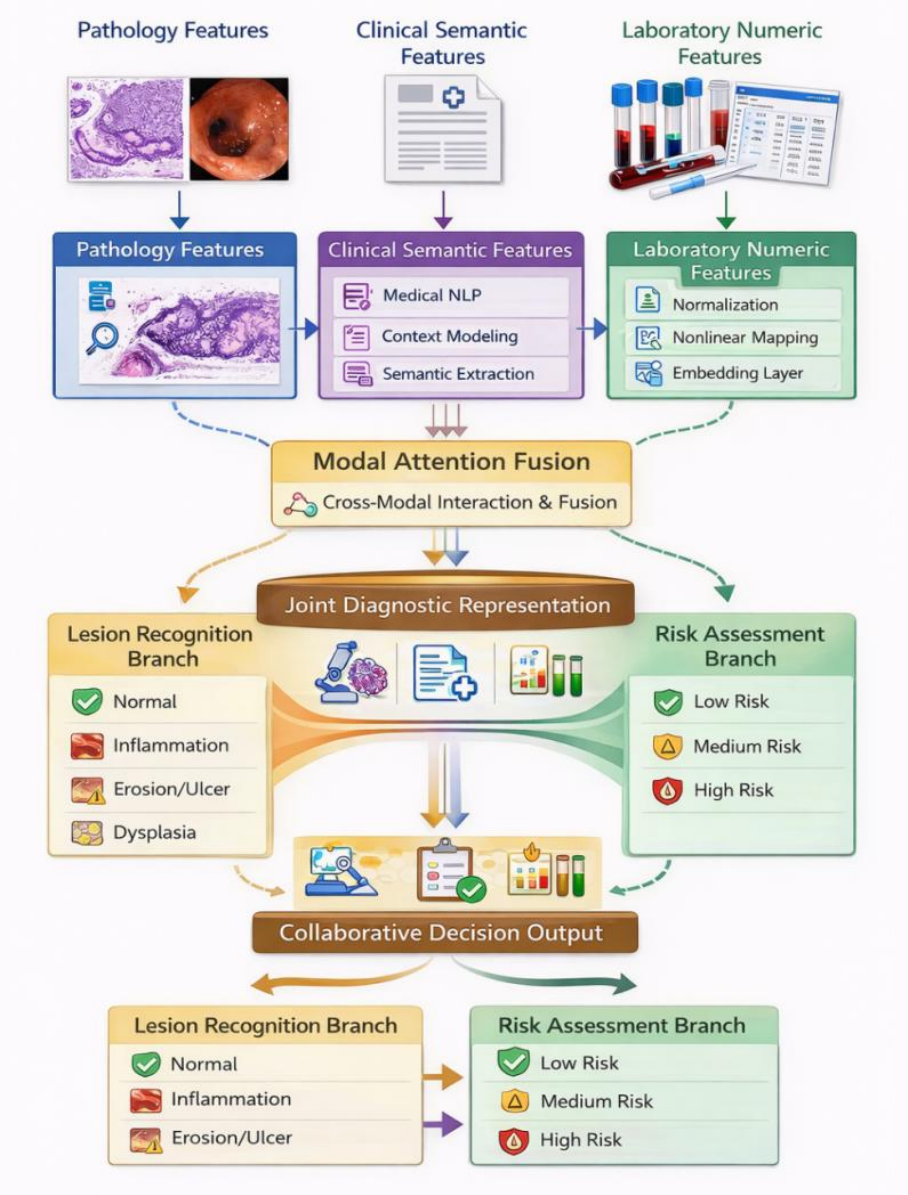

Figure 3: Diagram of collaborative decision-making mechanism for lesion recognition and risk discrimination

Let the shared features obtained for the ith case in the previous section be hiimg, hitxt, and hilab, respectively. In order to enhance the adaptive selection ability of the model for key modalities, this paper first performs fusion gate mapping on three types of features to obtain the modal interaction representation:

$$g_i = \mathrm{ReLU}\left(W_g\left[h_i^{img} || h_i^{txt} || h_i^{lab}\right] + b_g\right) \tag{13}$$

Here, || denotes the vector splicing operation, W_g and b_g are the fusion mapping parameters, and g_i is the case-level fusion context representation. This step can establish the joint relationship between different modalities in a unified space, so that the subsequent decision is no longer stopped at the simple superposition level, but based on the comprehensive characteristics after interaction.

Considering the different dependence degrees of different cases on the three types of

modalities, this paper further introduces the modal attention allocation mechanism to calculate the contribution weight of each modality in the current case:

$$\alpha_i^m=\frac{\exp(u_m^T g_i)}{\sum_{k\in\{img,txt,lab\}}\exp(u_k^T g_i)},\quad m\in\{img,txt,lab\} \tag{14}$$

Here, α_i^m represents the attention weight of the MTH modality in the ith case, and u_m is the learnable modality evaluation vector. Through this formula, the model can automatically adjust the participation intensity of images, texts and indicators in decision-making according to the specific performance of the case. For example, when the pathological section morphology is abnormal, the image modal weight will be increased. When persistent diarrhea, bleeding, or infection cues are present in the text, the effect of semantic modality is enhanced accordingly.

After the weights are obtained, the multimodal joint diagnostic representation of the case is constructed as follows:

$$f_i=\alpha_i^{img}h_i^{img}\oplus\alpha_i^{txt}h_i^{txt}\oplus\alpha_i^{lab}h_i^{lab}\oplus g_i \tag{15}$$

Here, $\oplus$ denotes the weighted stitching fusion operation, and f_i is the joint feature that is finally used for intelligent diagnosis.

At the decision level, this paper adopts a two-branch collaborative structure of lesion recognition and risk discrimination. The lesion recognition branch outputs the lesion category probability, which is used to distinguish normal mucosa, inflammatory injury, erosion ulcer, dysplasia and other states, and its expression is as follows:

$$p_i=\mathrm{Softmax}(W_d f_i+b_d) \tag{16}$$

Here, p_i represents the probability distribution of the lesion class, W_d and b_d are the classification branch parameters. At the same time, the risk discrimination branch combines the features and recognition results to output the risk intensity of the current case:

$$r_i=\sigma(W_r[f_i\|p_i]+b_r) \tag{17}$$

where $\sigma(\cdot)$ is the Sigmoid function, and r_i ranges from 0 to 1, with higher values indicating a higher likelihood of further deterioration or high-risk outcomes. The diagnosis probability and risk score are introduced into the discrimination process together, which helps to improve the sensitivity of the model to identify boundary cases and high-risk cases.

In order to realize the synchronous optimization of the two types of tasks, this paper uses the joint objective function for training:

$$L=L_{diag}+\lambda L_{risk}+\mu\|\Theta\|_2^2 \tag{18}$$

Here, L_{diag} is the lesion identification loss, L_{risk} is the risk discrimination loss, $\|\Theta\|_2^2$ is the parametric regularization term, and λ and μ are the balance coefficients. The objective function makes the model not only improve the classification accuracy, but also take into account the stable output of risk early warning and avoid the information skew caused by single task optimization.

Based on the above method, the model can extract complementary evidence from local pathological morphology, clinical semantic description and test value changes, and generate

the joint judgment result of lesion category and risk level through the collaborative decision-making mechanism. This design is suitable for application scenarios with complex case composition, rapid infection risk change, and high diagnostic timeliness requirements in post-earthquake environment, which can provide a more reliable decision-making basis for subsequent early warning scoring and stratified disposal.

3.4 Risk early warning scoring mechanism and early warning output strategy driven by diagnosis results

After the completion of lesion identification and risk discrimination, how to convert the diagnosis results into executable warning levels and disposal suggestions is the key link of the model application. Gastrointestinal mucosal lesions have the characteristics of rapid progression, complex causes and high risk of group transmission in post-earthquake environment. It is difficult to directly support clinical hierarchical treatment and emergency response by simply giving the lesion category or probability value. Therefore, based on the multimodal intelligent diagnosis results, this paper further constructs a risk early warning scoring mechanism driven by diagnosis results, which jointly maps the lesion severity, index abnormality, model risk output and diagnosis confidence information into continuous risk scores, and then generates early warning output through grade division and trigger rules. The risk warning level division and disposal suggestions are shown in Table 2.

Table 2: Risk warning level classification and disposal proposal Table

Warning Level	Score Range	Lesion Status Characteristics	Output Prompt	Management Recommendation	Response Time Limit
Level I Low Risk	0–34	The mucosa is basically stable or shows only mild inflammation, with slight fluctuations in indicators and relatively high model confidence	Routine Follow-up	Maintain oral rehydration and a clean diet, and recheck symptoms and basic indicators within 24–48 h	Within 48 h
Level II Moderate Risk	35–59	Inflammation is aggravated or there is a tendency toward focal erosion, with some infection-related indicators elevated and possible progression	Enhanced Monitoring	Intensify monitoring of body temperature, diarrhea frequency, and laboratory indicators; perform stool testing or infection screening when necessary	Within 24 h
Level III High Risk	60–79	Erosion or ulceration is evident, with increased risk of bleeding or persistent diarrhea, and abnormalities are concentrated across indicators	Priority Intervention	Initiate priority intervention, improve pathogen testing, endoscopic or pathological reassessment, and conduct rehydration and anti-infection evaluation	Within 6 h
Level IV Very High Risk	80–100	Severe mucosal injury or marked risk of dysplasia, with high probability of infection spread and complications	Emergency Alert	Trigger rapid multidisciplinary consultation and isolation management, and prioritize transfer to a facility with gastroenterology and infectious disease response capacity	Within 1 h

Let the lesion recognition branch output of the ith case be $p_i=[p_{i1},p_{i2},\ldots,p_{iK}]$, where different categories correspond to normal mucosa, inflammatory injury, erosion ulcer, dysplasia and other states. In order to reflect the clinical risk degree of different lesion categories, the lesion severity score is defined as follows:

$$d_i=\sum_{c=1}^{K}\gamma_c\, p_ic \tag{19}$$

Here, γ_c represents the severity coefficient corresponding to class c lesions, which increases progressively with the aggravation of the disease. This formula can transform the discrete lesion recognition results into continuous severity representations, making the diagnosis results more suitable for subsequent quantitative scoring.

Considering that gastrointestinal mucosal lesions are often accompanied by enhanced inflammation, aggravated infection and metabolic imbalance, this paper further introduces the detection index abnormal degree, which is used to describe the overall level of the body state deviation from the normal range. For the M core indicators included in the warning, the anomaly score is defined as follows:

$$a_i=\frac{1}{M}\sum_{j=1}^{M}\omega_j \min\left(\frac{|x_{ij}-r_j|}{u_j-r_j+\varepsilon},1\right) \tag{20}$$

Here, x_{ij} is the observed value of the JTH index in the ith case, r_j is the reference center value, u_j is the upper bound of the normal fluctuation of this index, and ω_j is the weight of the index. This formula suppresses extreme value interference by bounded normalization, so that different dimensional indicators can participate in scoring under a unified scale.

In order to avoid over-optimistic judgment in the low confidence state of the model, this paper introduces a diagnostic confidence correction term. Normalizing the information entropy of the probability distribution of lesion recognition, the following can be obtained.

$$c_i = 1-\frac{\sum_{c=1}^{K} p_ic \ln p_ic}{\ln K} \tag{21}$$

The closer c_i is to 1, the more concentrated the diagnosis distribution is and the more stable the model judgment is. When the c_i is low, it indicates that the case is in boundary status or multi-category mixed status and needs to be given higher attention in the warning score.

On this basis, the final risk score is constructed by integrating the lesion severity, index abnormality, risk discrimination value obtained in the previous section and confidence correction results:

$$R_i=100(\lambda_1 d_i+\lambda_2 a_i+\lambda_3 r_i+\lambda_4(1-c_i)) \tag{22}$$

where $\lambda_1,\lambda_2,\lambda_3$ and λ_4 are risk fusion coefficients, which satisfy:

$$\lambda_1+\lambda_2+\lambda_3+\lambda_4=1 \tag{23}$$

Thus, the continuous risk values can be uniformly mapped to the 0-100 interval, which is convenient for grading, threshold control and cross-case comparison.

Considering that the case status may change rapidly in a short time after the earthquake, this paper adopts the time smoothing strategy to dynamically update the risk score:

$$\widetilde{R}_i^{(t)}=\eta R_i^{(t)}+(1-\eta)\widetilde{R}_i^{(t-1)} \tag{24}$$

Here, $\widetilde{R}_i^{(t)}$ represents the smoothed risk value at time t, and η is the current observation weight. This treatment can weaken the false positives caused by a single detection fluctuation, while retaining the trend sensitivity when the condition continues to increase.

Finally, the system outputs the warning level according to the smoothed risk value:

$$G_i = \begin{cases} \text{Level I}, & 0 \le \widetilde{R}_i < 35 \\ \text{Level II}, & 35 \le \widetilde{R}_i < 60 \\ \text{Level III}, & 60 \le \widetilde{R}_i < 80 \\ \text{Level IV}, & \widetilde{R}_i \ge 80 \end{cases} \tag{25}$$

In the output strategy, the system not only gives the risk level, but also synchronously generates the lesion category, the key anomaly index, the main risk source and the suggested treatment path. For example, review recommendations and dietary hygiene intervention were exported for low and medium risk cases, key surveillance, pathogen screening and mucosal injury assessment were suggested for high risk cases, and prompt rapid consultation, isolation management and priority transport were triggered for extremely high risk cases.

4 Experimental results and analysis

4.1 Data set construction and experimental environment configuration

In order to verify the effectiveness of the multimodal intelligent diagnosis and risk warning method for gastrointestinal mucosal lesions proposed in this paper, a multimodal case dataset including pathological images, clinical texts and detection indicators is constructed, and the training and testing are completed in a unified experimental environment. The data samples covered four categories of normal mucosa, inflammatory injury, erosion/ulcer and dysplasia. The case-level samples were used as the basic statistical units, and the corresponding image blocks, medical record texts and laboratory test records were matched for each case. In order to ensure the comparability of training results, this paper divides the data into training set, validation set and test set according to the principle of class balance, and completes the parameter setting and model iteration on the same hardware platform. In addition, in order to correspond to the dynamic early warning strategy in Section 3.4, based on case-level static registration, this paper further sorted out the text update records and the change sequence of detection indicators in the continuous observation window after the treatment of some cases, and defined the time difference before the threshold of high risk judgment for the first time as the early warning time. It is used to evaluate the ability of the model to identify the trend of disease deterioration. Therefore, the experimental data can not only support the task of lesion category recognition, but also support the timeliness analysis of subsequent risk early warning, so as to ensure the consistency between model design, index setting and experimental verification. The data set composition and sample distribution are shown in Table 3.

Table 3: Data set composition and sample distribution table

Lesion Category	Number of Case Samples / cases	Number of Image Samples / images	Number of Text Records / records	Number of Indicator Records / sets	Training Set / cases	Validation Set / cases	Test Set / cases
Normal Mucosa	380	1140	380	380	266	57	57
Inflammatory Injury	520	1560	520	520	364	78	78
Erosion/Ulceration	430	1290	430	430	301	64	65
Dysplasia	290	870	290	290	203	43	44
Total	1620	4860	1620	1620	1134	242	244

As can be seen from Table 3, a total of 1620 case samples were included in the study, among which the inflammatory injury samples were the most, 520 cases, accounting for 32.10% of the total samples. There were 290 cases of dysplasia, accounting for 17.90%. From the overall structure, the four types of lesion samples have relatively stable data coverage, and the corresponding matching of images, text and index records can be realized at the case level, which provides a more complete data basis for subsequent cross-modal feature learning and joint diagnostic analysis. The histogram of the category distribution of the dataset is shown in Figure 4.

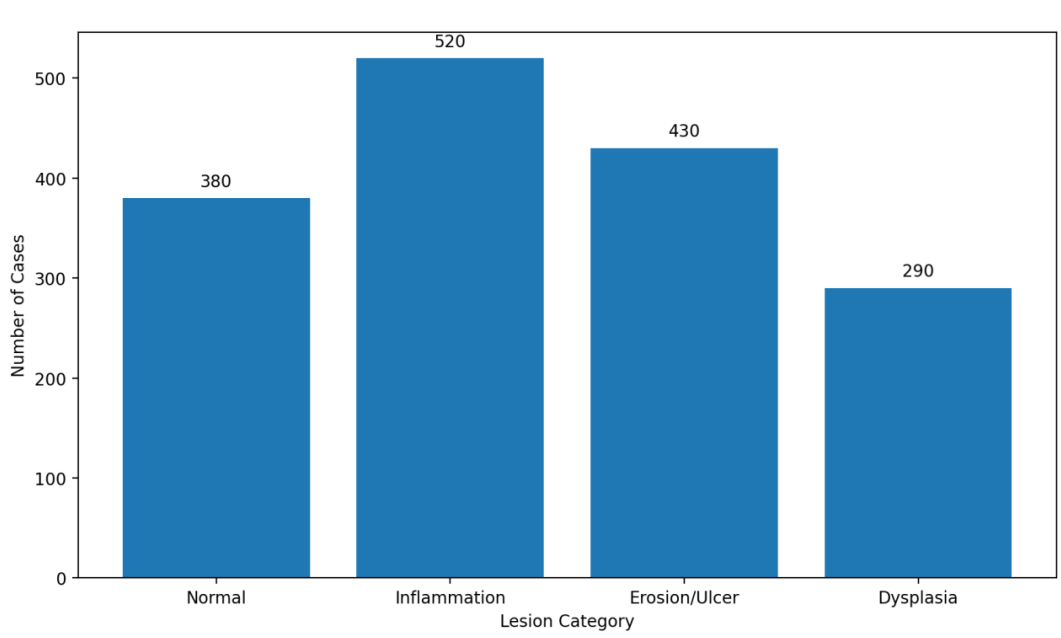

Figure 4: Histogram of class distribution of the dataset

It can be seen from Figure 4 that the number of inflammatory injury samples is the highest, the distribution of normal mucosa and erosion/ulcer samples is relatively close, and the number of dysplasia samples is small, but still maintains a basically usable category scale. In general, there is no excessive imbalance problem in the class distribution of the dataset, which helps to reduce the risk of bias for high-frequency categories in the model training process and improve the recognition stability of different lesion levels.

In terms of experimental environment, this paper uses Ubuntu 22.04 operating system,

Intel Xeon Silver 4314 processor, NVIDIA RTX 4090 graphics card and PyTorch 2.1 deep learning framework. The training rounds are set to 120, and the initial learning rate is set to 0.0001. The batch size is set to 16. On the whole, the current data scale, category structure and experimental environment configuration can meet the training and verification requirements of multimodal intelligent diagnosis and risk warning tasks.

4.2 Evaluation index system and design of comparative experiment scheme

In order to comprehensively evaluate the actual performance of the proposed method in the task of gastrointestinal mucosal lesion recognition and risk early warning, an evaluation index system is established from two levels of diagnostic performance and early warning performance. In the lesion recognition part, Accuracy, Precision, Recall, F1-score and AUC are selected as core indicators to measure the overall discrimination ability of the model for different lesion types, class balance performance and stability under threshold changes. Among them, Accuracy reflects the overall classification accuracy level, Precision and Recall correspond to the reliability of positive judgment and the ability to detect lesions, respectively. F1-score is used to comprehensively characterize the balance between precision and recall, and AUC is used to evaluate the comprehensive discrimination ability of the model under different discrimination thresholds. The risk Warning part is mainly evaluated by Warning Accuracy, Recall and Average Warning Lead Time, which are used to measure the correct degree of the model to identify the risk level, the detection ability of high-risk cases and the time advance amount of the warning output.

The comparison experiment uses unified data set division, unified training rounds and unified parameter search range to ensure the comparability of results between different methods. A total of five groups of comparison methods are set up in the experiment, which are single pathological image diagnosis method, single clinical text diagnosis method, pathological image and text fusion method, pathological image and detection index fusion method, and the full-modal fusion intelligent diagnosis and risk early warning method proposed in this paper. The single-modal method is used to verify the upper limit of the expression of a single information source, and the dual-modal method is used to investigate the gain effect of the local fusion strategy. The proposed method focuses on the comprehensive advantages of the collaborative modeling of images, text and detection indicators in lesion recognition and risk discrimination. In order to reduce the accidental interference, all experiments are run repeatedly under the same test set conditions and the average results are taken. At the same time, the contribution of key modules to the overall performance is further analyzed combined with subsequent ablation experiments.

4.3 Analysis of performance results of multimodal intelligent diagnosis

In order to verify the effectiveness of the proposed method in the recognition task of gastrointestinal mucosal lesions, the experiment selects a single pathological image method, a single clinical text method, a pathological image and text fusion method, a pathological image and detection index fusion method, and a full-modal fusion method proposed in this paper for comparative analysis. This paper focuses on the comprehensive performance of different methods in terms of Accuracy, F1-score and AUC. The comparison of diagnostic performance of different methods is shown in Figure 5.

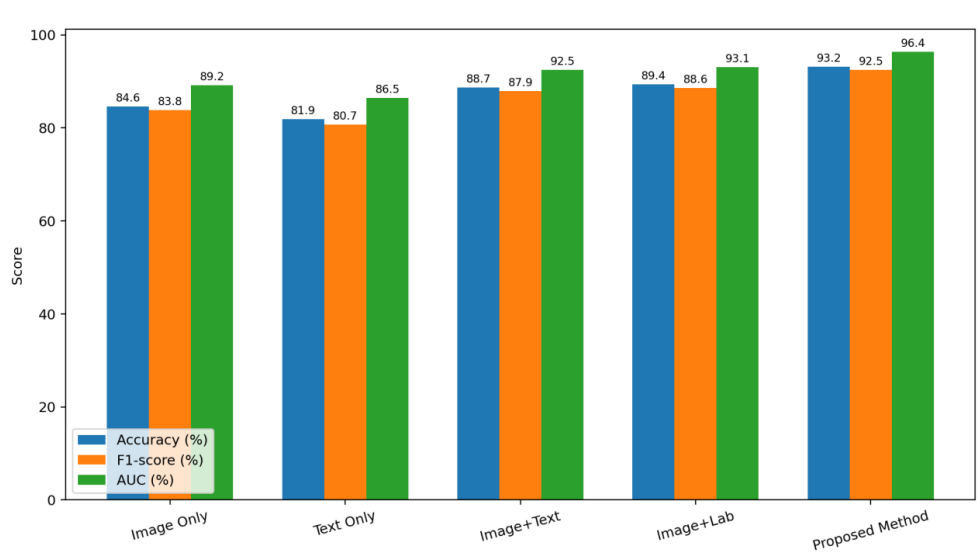

Figure 5: Bar charts comparing the diagnostic performance of different methods

It can be seen from Figure 5 that the proposed method achieves the best results on the three core indicators, where the Accuracy reaches 93.2%, the F1-score reaches 92.5%, and the AUC reaches 0.964, which is significantly better than that of the single pathological image method (84.6%, 83.8%, and 0.892). It is also higher than 81.9%, 80.7% and 0.865 of the single clinical text method. Compared with the two dual-modal methods, the proposed method still improves the Accuracy by 3.8 and 3.1 percentage points, and improves the F1-score by 4.6 and 3.9 percentage points, respectively, indicating that the collaborative fusion of image, text and detection indicators can more fully describe the morphological characteristics, clinical semantics and physiological changes of gastrointestinal mucosal lesions. Thus, the overall diagnostic accuracy and class discrimination ability are improved.

To further analyze the recognition effect of the proposed method on different lesion categories, the ROC curves of each lesion category are shown in Figure 6.

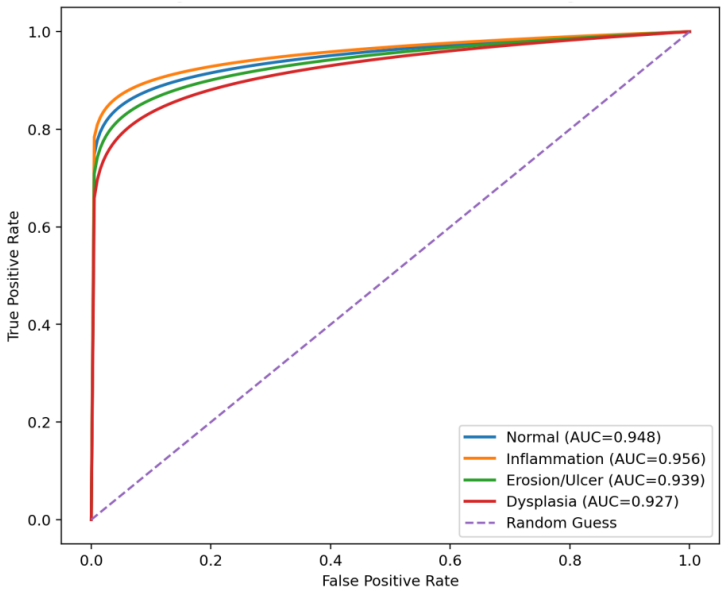

Figure 6: ROC curves for each lesion category

It can be seen from Figure 6 that the proposed method shows good discrimination ability in different lesion categories, and the AUCs of normal mucosa, inflammatory injury, erosion/ulcer and dysplasia are 0.948, 0.956, 0.939 and 0.927, respectively. The ROC curve of inflammatory lesion category was closest to the upper left corner, indicating that the recognition effect of this type of lesion was the most stable. The AUC of the dysplasia category is slightly lower, but still maintains above 0.92, indicating that the model still has a good ability to detect relatively complex and strong boundary lesions. On the whole, the multimodal intelligent diagnosis method proposed in this paper shows strong advantages in overall performance and classification recognition effect, which provides a reliable diagnosis basis for subsequent risk early warning analysis.

4.4 Analysis on identification effect and timeliness of risk early warning model

After the multimodal lesion recognition is completed, this paper further analyzes the recognition effect and timeliness of the risk warning module, focusing on the performance of the model in terms of early warning accuracy, recall rate and average early warning time under different risk levels. Since post-earthquake gastrointestinal mucosal lesions have the characteristics of uncertain progression, rapid transmission of infection and high risk of disease mutation, risk warning should not only ensure high recognition accuracy, but also find potential high-risk cases in advance as much as possible, so as to buy time for subsequent intervention. The comparison between accuracy and recall of risk warning is shown in Figure 7.

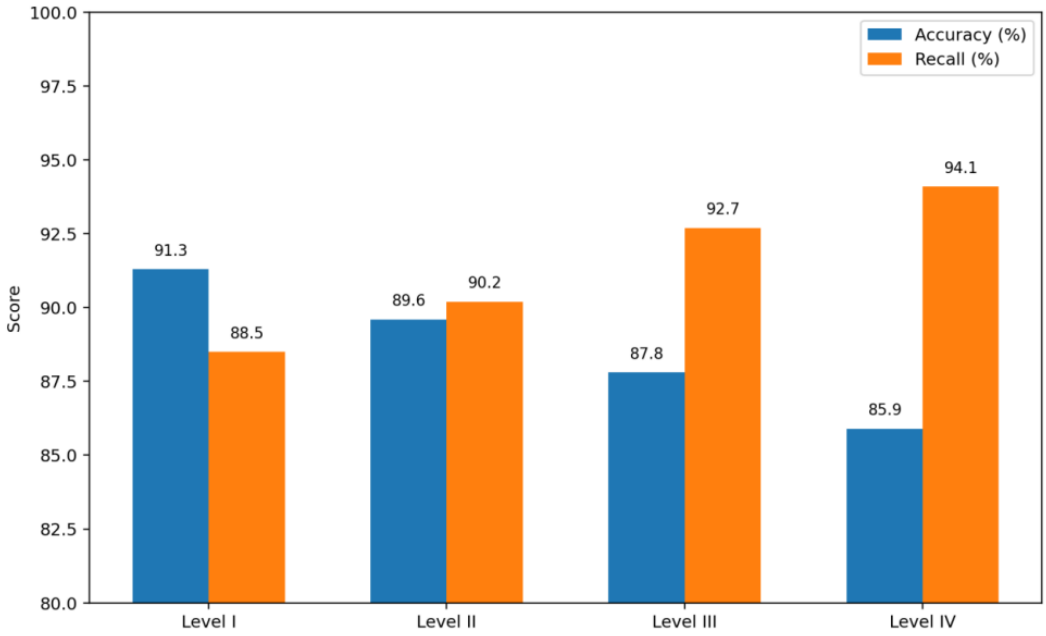

Figure 7: Comparison of accuracy and recall of risk warning

It can be seen from Figure 7 that the model maintains a high identification level under different risk levels. The precision and recall of Level I are 91.3% and 88.5%, respectively. With the increase of risk Level, the recall rate of the model gradually increased, reaching 92.7% and 94.1% at Level III and Level IV, respectively, indicating that the method in this paper has stronger detection ability for high-risk cases. Although the accuracy decreased slightly at the

high-risk Level, Level IV still maintained an accuracy of 85.9%, and the overall performance was relatively stable. This shows that the model can give priority to ensuring the identification sensitivity of severe and potentially worsening cases while controlling false positives, which is more in line with the actual needs of post-disaster emergency screening scenarios.Figure 8 shows the average warning lead time of different risk levels.

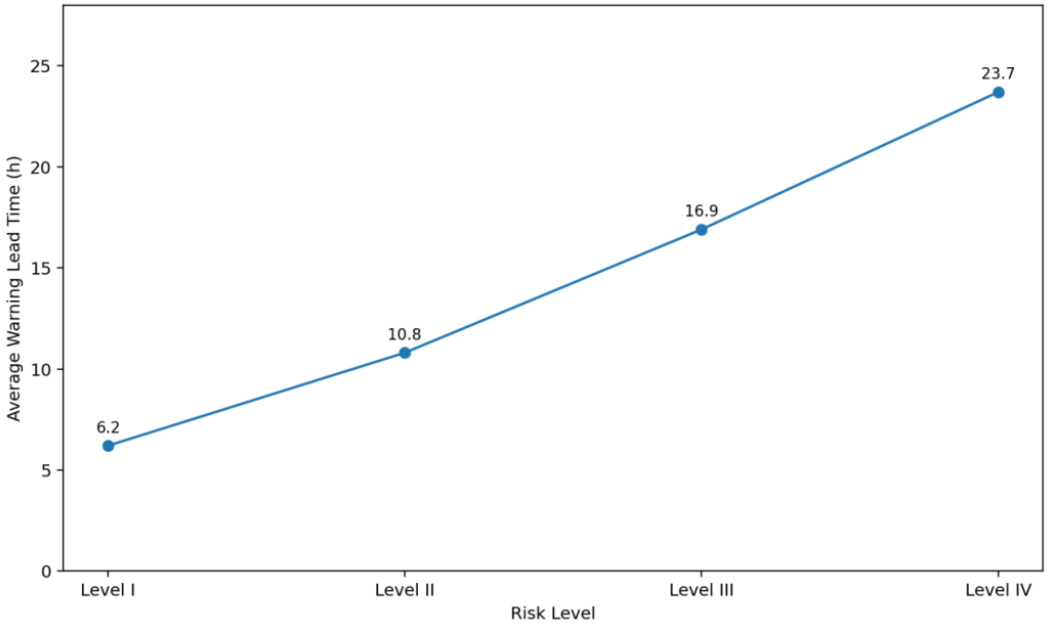

Figure 8: Line chart of average warning lead time for different risk levels

It can be seen from Figure 8 that the average warning lead time of the model increases significantly with the increase of the risk level. The average warning lead time of Level I cases was 6.2 hours, Level II cases increased to 10.8 hours, Level III cases and Level IV cases reached 16.9 hours and 23.7 hours, respectively. The results show that the proposed method not only has a good recognition effect for high-risk cases, but also can give early warning signals before the disease deteriorates further. In general, the risk warning module shows strong advantages in terms of recognition accuracy and timeliness, especially in the high-risk level, taking into account a high recall rate and a long advance, which can provide effective support for hierarchical monitoring and rapid intervention of post-earthquake gastrointestinal mucosal lesions.

4.5 Ablation experiments and model robustness analysis

In order to further verify the actual contribution of each key module in the proposed method and investigate the stability of the model under complex data perturbation conditions, this paper analyzes the ablation experiment and robustness test. The ablation experiment mainly removed the image modality, text modality, detection indicator modality, cross-modal alignment module and risk score module step by step, and compared the changes in Accuracy, F1-score, AUC and Warning Accuracy between the full model and each simplified model. It is used to identify the supporting effect of different components on the overall performance. At the same time, the robustness test further introduces different intensities of noise interference and different proportions of data missing into the test set to evaluate the adaptability of the model to incomplete data and low-quality data in complex post-earthquake scenes. The results of ablation experiments are shown in Table 4.

Table 4: Comparison table of ablation experiment results

Model Configuration	Accuracy / %	F1-score / %	AUC	Warning Accuracy / %
Full Model	93.2	92.5	0.964	90.8
Without Image Modality	88.6	87.9	0.921	85.1
Without Text Modality	89.1	88.4	0.927	85.9
Without Laboratory Indicator Modality	89.7	88.9	0.933	86.6
Without Cross-modal Alignment Module	90.3	89.8	0.938	87.4
Without Risk Scoring Module	91.0	90.4	0.946	84.7

It can be seen from Table 4 that the full model achieves optimal results in all indicators. Among them, the Accuracy decreased to 88.6% and F1-score decreased to 87.9% after removing the image modality, indicating that pathological morphological information is still the core basis for the recognition of gastrointestinal mucosal lesions. After removing the text modality and the detection index modality, the model performance also decreases to varying degrees, indicating that clinical semantic information and numerical indicators play an important complementary role in the discrimination of complex cases. After removing the cross-modal alignment module, the AUC decreases to 0.938, indicating that the multimodal consistency modeling has an obvious effect on improving the stability of the overall classification boundary. After removing the risk scoring module, the Warning Accuracy is reduced to 84.7%, which further indicates that this module improves the quality of warning output most directly.The model performance variation under noise intensity or missing ratio is shown in Figure 9.

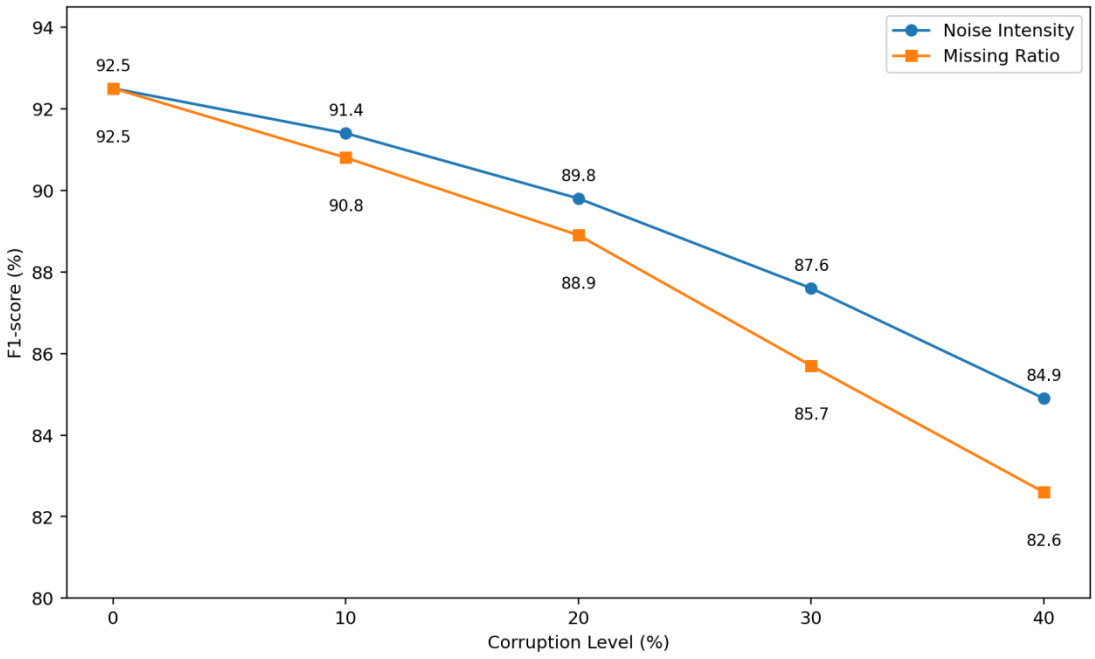

Figure 9: Line plot of model performance change under noise intensity or missing ratio

As can be seen from Figure 9, as the noise intensity and the proportion of missing data increase, the F1-score of the model shows a downward trend as a whole, but the change

process is relatively stable. When the disturbance level increased from 0% to 40%, the F1-score decreased from 92.5% to 84.9% in the noise condition and from 92.5% to 82.6% in the absence condition. In contrast, the increase of missing ratio has a more obvious impact on model performance, indicating that data integrity is more critical for multimodal collaborative discrimination. In general, the proposed method can still maintain high recognition performance in the medium disturbance range, reflecting good robustness and adaptability to complex environments after disasters.

5 Discussion

The results of this study show that the intelligent diagnosis and risk early warning method driven by multimodal pathological data can better improve the recognition accuracy of gastrointestinal mucosal lesions, and show a strong time-efficiency advantage in the screening of high-risk cases. This result indicates that under the situation of limited post-earthquake medical conditions, rapid increase of cases, and continuous rise of manual interpretation pressure, collaborative analysis of pathological images, clinical texts and detection indicators can better reflect the real state of lesions than a single information source. Especially for inflammatory injury, erosive ulcer and early high-risk lesions, multimodal fusion not only improves the misjudgment problem caused by fuzzy classification boundaries, but also enhances the stability of risk stratification judgment, indicating that this method has good application value for post-disaster emergency screening.

Post-earthquake environmental mutation, survival pressure and continuous uncertainty can easily induce acute stress disorder, and then evolve into chronic anxiety. Accurate identification and modeling of anxiety disorders is an important technical means for early screening of high-risk groups and avoiding comprehensive psychological collapse. From the overall perspective of post-disaster medical relief, psychological crisis and physical disease are not isolated from each other. Anxiety, sleep disorder and stress imbalance will further weaken the body's immune regulation ability and life behavior control ability, making the affected people more prone to eating disorders, dehydration, infection spread and basic disease decompensation. Therefore, the construction of post-disaster intelligent recognition system should not only stay at the level of single disease diagnosis, but also develop towards the linkage direction of mental state recognition, early warning of digestive system damage and infection transmission monitoring. This paper focuses on gastrointestinal mucosal lesions, but its method framework also suggests that more comprehensive multi-source medical perception and collaborative research mechanism should be established for the identification of high-risk populations after disasters.

After the earthquake, facilities are damaged, water and food are easily contaminated, bacteria directly attack the gastrointestinal mucosa, and the risk of infectious diseases rises sharply. Constructing an accurate model of gastrointestinal mucosal lesions is the core key to realize early diagnosis and treatment and block the deterioration of the epidemic. The experimental results show that the multimodal fusion method is superior to the single-modal and dual-modal methods in terms of diagnostic accuracy, risk recognition effect and warning advance, which indicates that gastrointestinal mucosal lesions can not be fully judged by a single image observation, but need to be comprehensively evaluated by combining clinical semantic clues and changes in test indicators. Of course, there are still some limitations in this paper, such as relatively concentrated data source scenarios, performance degradation under extreme missing conditions, and early warning threshold setting need to be verified with more clinical follow-up. Follow-up studies can further introduce time-series monitoring data, post-disaster environmental exposure information, and regional infection epidemic

characteristics to improve the generalization ability and practical value of the model in real emergency treatment scenarios.

6 Conclusion

Focusing on the problems of difficult identification of post-earthquake gastrointestinal mucosal lesions, rapid risk evolution, and heavy burden of manual interpretation, this paper proposes an intelligent diagnosis and risk early warning method driven by multimodal pathological data. Starting from pathological images, clinical texts and detection indicators, data standardization processing, cross-modal coding and alignment are completed. Combined with modal attention fusion, dual-branch collaborative discrimination and risk scoring mechanism, the integration of lesion recognition and early warning output is realized. Experiments show that the complete model achieves 93.2% Accuracy, 92.5% F1-score and 0.964 AUC on 1620 samples, and the warning accuracy reaches 90.8%. Under the condition of 40% noise and missing, F1-score still maintains 84.9% and 82.6%, showing good robustness. This method can provide effective support for post-earthquake high-risk case screening, clinical rapid decision making and infection risk interruption, and needs to be further improved by combining time series monitoring and real emergency scenarios in the future.

References

[1] Calderaro J, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. Gut. 2021;70(6):1183-1193. doi:10.1136/gutjnl-2020-322880.

[2] Berbís MA, Aneiros-Fernández J, Mendoza Olivares FJ, Nava E, Luna A. Role of artificial intelligence in multidisciplinary imaging diagnosis of gastrointestinal diseases. World J Gastroenterol. 2021;27(27):4395-4412. doi:10.3748/wjg.v27.i27.4395.

[3] Suzuki H, Tokai Y, Yoshio T, Tada T. Artificial intelligence for cancer detection of the upper gastrointestinal tract. Dig Endosc. 2021;33(2):254-262. doi:10.1111/den.13897.

[4] Correia FP, Lourenço LC. Artificial intelligence application in diagnostic gastrointestinal endoscopy - Deus ex machina? World J Gastroenterol. 2021;27(32): 5351-5361. doi:10.3748/wjg.v27.i32.5351.

[5] Ali H, Muzammil MA, Dahiya DS, et al. Artificial intelligence in gastrointestinal endoscopy: a comprehensive review. Ann Gastroenterol. 2024;37(2):133-141. doi:10.20524/aog.2024.0861.

[6] Iacucci M, Santacroce G, Zammarchi I, et al. Artificial intelligence and endo-histo-omics: new dimensions of precision endoscopy and histology in inflammatory bowel disease. Lancet Gastroenterol Hepatol. 2024;9(8):758-772. doi:10.1016/S2468-1253(24)00053-0.

[7] Ueyama H, Kato Y, Akazawa Y, et al. Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. J Gastroenterol Hepatol. 2021;36(2):482-489. doi:10.1111/jgh.15190.

[8] He X, Wu L, Dong Z, et al. Real-time use of artificial intelligence for diagnosing early gastric cancer by magnifying image-enhanced endoscopy: a multicenter diagnostic study (with videos). Gastrointest Endosc. 2022;95(4):671-678.e4. doi:10.1016/j.gie.2021.11.040.

[9] Horiuchi Y, Hirasawa T, Fujisaki J. Application of artificial intelligence for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. Clin Endosc. 2024;57(1):11-17. doi:10.5946/ce.2023.173.

[10] Jin C, Jiang Y, Yu H, et al. Deep learning analysis of the primary tumour and the prediction of lymph node metastases in gastric cancer. Br J Surg. 2021;108(5):542-549. doi:10.1002/bjs.11928.

[11] Huang B, Tian S, Zhan N, et al. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. EBioMedicine. 2021;73:103631. doi:10.1016/j.ebiom.2021.103631.

[12] Sirinukunwattana K, Domingo E, Richman S, et al. Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. Gut. 2021;70(3):544-554. doi:10.1136/gutjnl-2019-319866.

[13] Shimada Y, Okuda S, Watanabe Y, et al. Histopathological characteristics and artificial intelligence for predicting tumor mutational burden-high colorectal cancer. J Gastroenterol. 2021;56(6):547-559. doi:10.1007/s00535-021-01789-w.

[14] Kudo SE, Ichimasa K, Villard B, et al. Artificial Intelligence System to Determine Risk of T1 Colorectal Cancer Metastasis to Lymph Node. Gastroenterology. 2021; 160(4):1075-1084.e2. doi:10.1053/j.gastro.2020.09.027.

[15] Neumann H, Kreft A, Sivanathan V, Rahman F, Galle PR. Evaluation of novel LCI CAD EYE system for real time detection of colon polyps. PLoS One. 2021; 16(8):e0255955. doi:10.1371/journal.pone.0255955.

[16] Sun C, Li B, Wei G, et al. Deep learning with whole slide images can improve the prognostic risk stratification with stage III colorectal cancer. Comput Methods Programs Biomed. 2022;221:106914. doi:10.1016/j.cmpb.2022.106914.

[17] Huang K, Lin B, Liu J, et al. Predicting colorectal cancer tumor mutational burden from histopathological images and clinical information using multi-modal deep learning. Bioinformatics. 2022;38(22):5108-5115. doi:10.1093/bioinformatics/btac641.

[18] Xu Y, Jiang L, Chen W, et al. Computer-aided detection and prognosis of colorectal cancer on whole slide images using dual resolution deep learning. J Cancer Res Clin Oncol. 2023;149(1):91-101. doi:10.1007/s00432-022-04435-x.

[19] Lim Y, Choi S, Oh HJ, et al. Artificial intelligence-powered spatial analysis of tumor-infiltrating lymphocytes for prediction of prognosis in resected colon cancer. npj Precis Oncol. 2023;7(1):124. doi:10.1038/s41698-023-00470-0.

[20] Hüneburg R, Bucksch K, Schmeißer F, et al. Real-time use of artificial intelligence (CADEYE) in colorectal cancer surveillance of patients with Lynch syndrome-A

randomized controlled pilot trial (CADLY). United European Gastroenterol J. 2023;11(1):60-68. doi:10.1002/ueg2.12354.

[21] Li JW, Wu CCH, Lee JWJ, et al. Real-World Validation of a Computer-Aided Diagnosis System for Prediction of Polyp Histology in Colonoscopy: A Prospective Multicenter Study. Am J Gastroenterol. 2023;118(8):1353-1364. doi:10.14309/ajg.0000000000002282.

[22] Repici A, Spadaccini M, Antonelli G, et al. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. Gut. 2022;71(4):757-765. doi:10.1136/gutjnl-2021-324471.

[23] Yao L, Li X, Wu Z, et al. Effect of artificial intelligence on novice-performed colonoscopy: a multicenter randomized controlled tandem study. Gastrointest Endosc. 2024;99(1):91-99.e9. doi:10.1016/j.gie.2023.07.044.

[24] Ortiz O, Macaron C, Zubiaurre L, et al.; TIMELY study group. An artificial intelligence-assisted system versus white light endoscopy alone for adenoma detection in individuals with Lynch syndrome (TIMELY): an international, multicentre, randomised controlled trial. Lancet Gastroenterol Hepatol. 2024;9(9):802-810. doi:10.1016/S2468-1253(24)00187-0.

[25] Jiang X, Hoffmeister M, Brenner H, et al. End-to-end prognostication in colorectal cancer by deep learning: a retrospective, multicentre study. Lancet Digit Health. 2024;6(1):e33-e43. doi:10.1016/S2589-7500(23)00208-X.

[26] Xu Y, Guo J, Yang N, et al. Predicting rectal cancer prognosis from histopathological images and clinical information using multi-modal deep learning. Front Oncol. 2024;14:1353446. doi:10.3389/fonc.2024.1353446.