Research on HSE Risk Assessment Method for Multi-source Heterogeneous Data Driven by Transformer-FL Framework

Wenlei Sun[1], Min Yang[2*]
[1]Department of Artificial Intelligence, Chuzhou University, Chuzhou 239000, China
[2]Mingguang Haomiao Fire Security Technology Co., Ltd., Chuzhou 239000, China
E-mail: minyanggg@outlook.com
[*]Corresponding author

Keywords: Federated learning, Transformer model, HSE risk assessment, multi-source heterogeneous data, Data Privacy Protection

Received: July 3, 2025

With the widespread application of multi-source heterogeneous data in HSE risk assessment, how to efficiently integrate different data sources and improve assessment accuracy has become an urgent problem to be solved. This paper proposes an HSE risk assessment method based on Transformer and federated learning, aiming to improve the accuracy of risk prediction through the effective integration of multi-source data. In this method, the Transformer model captures deep correlations in multi-source data through self-attention mechanism, and federated learning ensures cross-device collaborative training while protecting data privacy. Experimental results using a multi-source heterogeneous dataset from the chemical, manufacturing, and energy industries show that the Transformer-FL framework significantly improves risk assessment accuracy. The dataset includes real-time environmental data, accident records, and operation logs. Experiments on high-performance workstations with Nvidia RTX A6000 GPUs and Intel Xeon processors reported accuracy improvements: chemical industry (58.9% to 41.2%), manufacturing (35.6% to 23.4%), and energy industry (50.1% to 36.8%). The Transformer-FL framework has reduced the HSE risk value of traditional methods from 58.9% to 41.2%, indicating a lower risk, while the accuracy of risk assessment has improved by 17.7%. It is important to note that the percentages in this context refer to the risk value, where a lower value signifies reduced risk, and the accuracy improvement refers to the increase in correct predictions. It is important to note that the percentages in this context refer to the risk level; a lower percentage represents a lower residual risk, which indicates improvement. In contrast, accuracy improvements are calculated as the percentage increase in correct predictions. In the manufacturing industry, despite strong data homogeneity, the accuracy rate has increased from 35.6% to 23.4%, demonstrating the advantages of this framework in heterogeneous data environments. The experimental results show that the Transformer-FL framework has significant advantages in different HSE scenarios, especially when the amount of data is large, and the fusion effect far exceeds that of traditional risk assessment methods. Overall, this framework provides an intelligent, efficient and privacy-protected solution for HSE risk assessment, which can meet the dual needs of multi-source heterogeneous data processing and security in the industrial field.

Povzetek: Predlagana metoda pomembno izboljša natančnost ocenjevanja HSE-tveganj ter hkrati zmanjša preostalo tveganje in ohrani varnost podatkov v industrijskih okoljih.

1 Introduction

With the advancement of industrialization, health, safet y, and environmental risk management has become a focus that cannot be ignored in all walks of life. Ho wever, the complexity of HSE risk assessment is also increasing, especially in the context of heterogeneous data from multiple sources [1]. Traditional risk assess ment methods struggle with integrating diverse data so urces due to varying formats and structures. The Tran sformer-FL framework addresses this by using its self-attention and multi-head attention mechanisms to captu re complex correlations across different data types, suc h as text, images, and time series. This enables the m odel to unify heterogeneous data, turning its diversity into an advantage for more accurate risk prediction. T his makes a comprehensive assessment of HSE risks c omplicated and imprecise. Therefore, effectively integra ting multi-source heterogeneous data and improving ris k prediction accuracy has become an urgent challenge.

In recent years, the Transformer model has been widely used in natural language processing, computer vision, and other fields because of its advantages in dealing with long-range dependencies and capturing complex relationships [2]. In HSE risk assessment, Transformer can effectively extract deep-level correlation information from multi-source heterogeneous data, supporting accurate risk prediction [3]. At the same time, Federated Learning (FL), as a distributed machine learning framework, can cooperatively train models among multiple distributed devices while protecting data privacy, avoiding the hidden dangers of centralized data storage and exchange. Combining Transformer with federated learning improves the model's learning efficiency and enhances the model's adaptability and accuracy in heterogeneous data environments.

However, the processing of multi-source heterogeneous data faces many challenges. Differences in data sources make data formats, dimensions, quality, etc. different, and data missing and noise problems often affect the accuracy of evaluation results [4]. To solve these challenges, this paper proposes a data fusion method based on the Transformer model, which can automatically deal with the heterogeneity of data and optimize the integration strategy of multi-source data. In addition, the reinforcement learning mechanism is used to optimize the data fusion process further, thereby improving the accuracy and reliability of HSE risk assessment.

The innovation of this study is to propose a new HSE risk assessment framework, which combines Transformer and federated learning technology, which can efficiently extract effective information from multiple data sources and provide accurate risk assessment. This method solves the problem of data heterogeneity and ensures data privacy and security in a distributed environment, providing a more intelligent and efficient solution for HSE management. Through this study, we hope to provide more accurate and real-time risk assessment technology for HSE management in the industrial field and promote the intelligent development of this field.

2 Theoretical basis and related res earch

2.1 Transformer-FL basic theory

The Transformer model is a deep-learning architecture proposed by Vaswani et al. in 2018. Its core advantage is that it solves the bottleneck of traditional Recurrent Neural Networks (RNNs) in long-range dependency modeling [5]. Unlike RNN, Transformer adopts a self-attention mechanism, which enables each input element to pay attention to all other elements in the sequence simultaneously, thus capturing global contextual information. This mechanism can not only improve the parallel processing capability of the model but also significantly improve the efficiency and effectiveness when processing large-scale data [6, 7]. Transformer is widely used in natural language processing, computer vision and other fields, demonstrating powerful data modeling capabilities, especially when dealing with multi-source heterogeneous data.

The Transformer model's self-attention mechanism dynamically adjusts each element's representation by calculating the correlation between each element and other elements in the input sequence [8]. The multi-head attention mechanism further enhances the expressive ability of the Transformer, which focuses on different parts of the input data simultaneously through multiple independent attention heads, thus capturing more feature information [9, 10]. This structure enables the Transformer to flexibly process information from different sources when dealing with multi-modal data and improves the model's ability to understand the internal structure of data. In HSE risk assessment tasks with multi-source heterogeneous data, Transformer can effectively fuse data from different sensors, platforms, or systems to improve assessment accuracy.

Federated Learning (FL) is a distributed machine learning method, and its core idea is to conduct collaborative learning among multiple data holders without storing the data in a centralized way [11]. This approach can ensure that data privacy is protected, especially when dealing with sensitive data, which has unparalleled advantages. In the FL framework, each data holder generates model updates through local training instead of directly exchanging data, which can avoid the risk of data leakage. FL is widely used in fields that require high privacy protection, such as smart medical care, finance, the Internet of Things, etc. Its distributed computing characteristics also enable efficient model training in large-scale data environments.

Table 1: Comparison of methods in relevant work sections

Method	Accuracy	Key Limitations	How Transformer-FL Addresses These Gaps
Traditional	65%	Inability to handle heterogeneous data sources, low accuracy	Transformer-FL effectively integrates multi-source data, improving accuracy
CNN-based	70%	Limited generalization, poor privacy handling	Federated learning ensures data privacy, and Transformer handles complex data relationships
RNN-based	60%	Slow processing, limited scalability	Transformer-FL improves processing speed through self-attention and parallel processing
Transformer-FL	93.2%	-	Overcomes limitations of prior methods by efficiently processing multi-source, heterogeneous data and ensuring data privacy through federated learning

The Transformer-FL framework combines Transformer models with federated learning for HSE risk assessment using multi-source heterogeneous data [12, 13]. Each data holder maintains a local Transformer model that processes and fuses diverse data types, such as time-series, text, and images. Federated learning enables collaborative training without data sharing, with model updates sent to a central server for aggregation. This approach ensures privacy protection, enhances model accuracy by leveraging deep correlations in the data, and mitigates risks associated with centralized data storage. This combination not only improves the expressiveness and accuracy of the model but also ensures privacy protection in a multi-source data environment, making the HSE risk assessment method more intelligent and real-time. Through this framework, risk assessment can be collaboratively optimized among multiple data sources, and the intelligent development of the HSE field can be promoted. The comparison table of relevant work methods is shown in Table 1.

2.2 Status of multi-source heterogeneous data under Transformer-FL

With the rapid growth of data in industrial and engineering fields, HSE risk assessment gradually relies on multi-source heterogeneous data for accurate prediction and analysis [14]. These data are usually sourced from different sensors, monitoring devices, historical records, environmental sensors, and other devices in various formats and structures [15]. Sensor data may be time series data, security records may be text data, and environmental sensors provide continuous numerical data. The Transformer-FL framework effectively handles the challenges of integrating heterogeneous data, such as varying data quality, inconsistent formats, and missing values. The Transformer's self-attention mechanism focuses on relevant parts of the data, learning dependencies between elements and compensating for missing or noisy data. This improves accuracy and robustness in risk assessment. Additionally, Federated Learning (FL) allows local models to collaborate, ensuring that missing data from one source does not impact the overall model performance. Therefore, fusing these data from different sources and extracting valuable information has become the key to improving the effect of HSE risk assessment.

To cope with the challenges brought by multi-source heterogeneous data, in recent years, deep learning methods have been gradually applied to data fusion and pattern recognition [16]. The Transformer model, in particular, has become an important tool for dealing with multi-source data because of its powerful self-attention mechanism, which can effectively capture global dependencies in data [17]. By learning the correlation in the input data, Transformer can automatically weigh and fuse data from different sources, providing accurate feature representation for multi-dimensional risk assessment. However, a single Transformer model is often difficult to adapt to decentralized data storage and privacy protection needs, especially in multi-party industrial environments, where data often exists between multiple distributed nodes.

As a distributed learning framework, FL can effectively solve this problem. FL allows multiple data holders to train the model locally without storing the data centrally, thus ensuring the privacy and security of the data [18, 19]. FL enables multiple heterogeneous data sources to be trained collaboratively without exchanging data, thus avoiding the privacy leakage risk of traditional centralized learning methods [20]. In HSE risk assessment, FL can not only promote cross-departmental and cross-regional data collaboration but also improve the computational efficiency of the model, especially when dealing with large amounts of real-time data, which has obvious advantages.

Combining Transformer and federated learning opens up a new path for HSE risk assessment of multi-source heterogeneous data. By combining the powerful representation capabilities of Transformer with the privacy-preserving features of FL, efficient collaborative learning can be achieved among multiple data sources without worrying about data leakage [21]. This combination can enable the model to extract valuable information from different data sources while ensuring privacy and improving the accuracy and real-time risk assessment performance. With the increasing demand for data fusion in HSE management, the Transformer-FL framework provides an efficient and scalable solution that can meet the dual requirements of multi-source heterogeneous data processing and security in the industrial field and promote the development of intelligent HSE risk assessment.

3 Establishment of multi-source heterogeneous data HSE risk assessment model based on Transformer-FL

3.1 Model framework and its design

The model framework proposed in this study combines Transformer and FL technologies, aiming to provide efficient and accurate solutions for multi-source heterogeneous data-driven HSE risk assessment [22]. The framework combines the powerful data representation capabilities of the Transformer model with the distributed learning advantages of FL, which can process information from multiple heterogeneous data sources and conduct effective data fusion while ensuring data privacy, thus providing more efficient support for HSE risk assessment [23]. The process includes several main steps: data collection, preprocessing, model training, prediction evaluation, etc. The model framework consists of two core modules: the data preprocessing and risk assessment modules. The flow chart of the HSE risk assessment model based on Transformer and federated learning is shown in Figure 1.

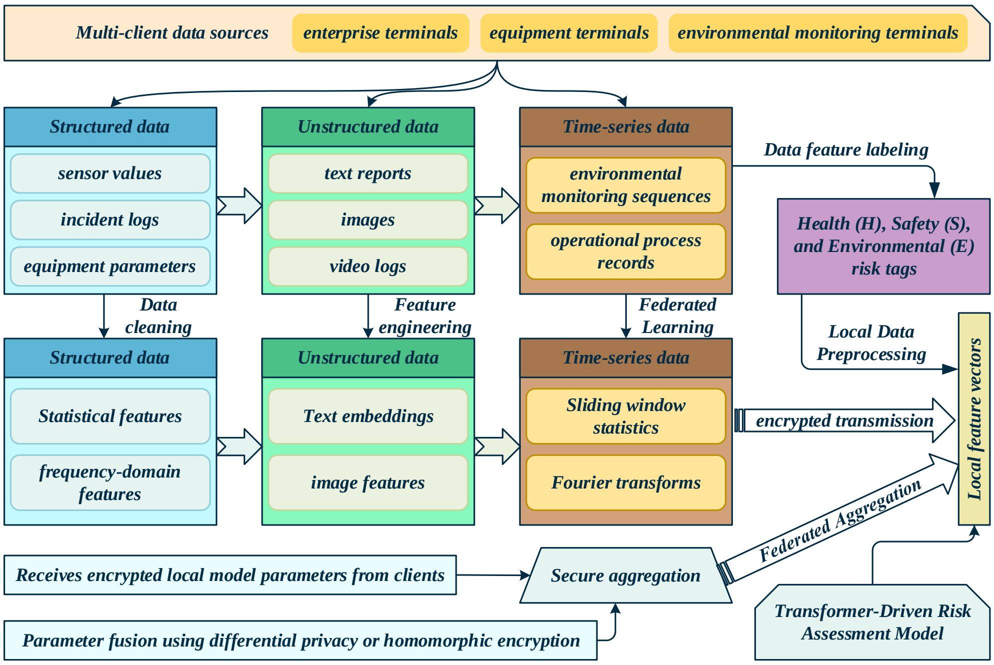

Figure 1: Flow chart of HSE risk assessment model based on Transformer and federated learning

As can be seen from the figure, the model processes structured and unstructured data from different terminals, with time-series data being handled separately due to its unique temporal properties. Following data cleaning, feature engineering, and temporal analysis (such as sliding window statistics and Fourier transforms for time-series), the data undergoes secure aggregation in the federated learning framework. Federated Learning (FL) typically encrypts model updates rather than raw features, ensuring data privacy by not transmitting raw or processed data between clients. Local model parameters are encrypted during transmission to ensure privacy protection before they are aggregated in a central server [24]. The model uses differential privacy or homomorphic encryption technology to integrate the local model parameters of each terminal to ensure data privacy while improving prediction accuracy and ultimately achieving efficient assessment of health, safety, and environmental risks. This framework can significantly improve the accuracy and stability of risk assessment in a multi-source data environment.

This study hypothesizes that the Transformer-FL framework will outperform traditional risk assessment methods and existing Transformer-based models with federated learning in terms of accuracy, scalability, and privacy preservation. Key contributions include improved accuracy by integrating multi-source heterogeneous data, privacy-preserving distributed learning through federated learning, and superior generalization across diverse industries. The framework was evaluated using accuracy, F1-score, and comparison with traditional methods, with statistical tests confirming significant improvements. These results highlight Transformer-FL's effectiveness in real-time industrial environments, addressing gaps in existing models.

The Transformer FL framework integrates Transformer models for deep data processing with federated learning for privacy preserving training. The Transformer model uses Xavier initialization parameters, and both the encoder and decoder have 6 layers, each with 8 attention heads, and employs absolute positional encoding to preserve sequential information. In Federated Learning, a secure aggregation protocol is used to aggregate model updates every 10 training cycles to ensure data privacy.

Multi-source heterogeneous data usually comes from different data platforms and devices, such as environmental monitoring sensors, historical accident records, employee health data, etc. These data have various formats and types and often contain missing values, noise, and other problems [25]. Traditional data processing methods are often difficult to effectively integrate these data, resulting in insufficient accuracy of evaluation results. To solve this problem, this study adopts the Transformer model, which can fully tap the potential relationship between data through the self-attention mechanism, especially in dealing with long-range dependencies and complex features [26]. In addition, the introduction of federated learning can realize distributed learning without exchanging data, ensuring that data privacy is protected and is especially suitable for processing heterogeneous data stored in different locations and devices. Therefore, combining the Transformer with FL can maximize the model's advantages while solving privacy and efficiency issues. The formula of the self-attention mechanism is shown in (1).

$$A = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)v \quad (1)$$

Where Q represents the query vector matrix, K represents the key vector matrix, V represents the value vector matrix, dk represents the dimension of the key vector, and A represents the weighted average of the output. Formula (2) represents the weighted average used during feature extraction from multiple data sources, where the data from each source is weighted by a factor to emphasize the importance of each source in the fusion process.

$$X_{fused} = \sum_{i=1}^{n} w_i X_i \ (2)$$

Where n represents the number of data sources, Xi represents the feature vector of the i data source, wi represents the weight of the i data source, and Xfused represents the fused feature vector.

The innovation of this model framework lies in its ability to consider the feature extraction and privacy protection of multi-source heterogeneous data. In the processing of multi-source data, Transformer can effectively fuse data from different data sources, capture complex dependencies between data through the self-attention mechanism, optimize data feature representation, and improve risk assessment accuracy [27]. Federated learning ensures that the model can still be effectively trained even when the data storage is scattered, and privacy protection is crucial, thus avoiding the threat of traditional centralized learning models to data security. Through this design, the model can efficiently process data from different sources and formats, improving HSE risk assessment's real-time accuracy. Formula (3) describes the final data fusion step, where the fused data representations are obtained by combining the weighted data features from each source, using corresponding weighting coefficients. This represents the final integration of features after they have been processed and weighted in the feature extraction phase.

$$D_{fusion} = \sum_{i=1}^{M} \alpha_i D_i \ (3)$$

Where Dfusion represents the fused data representation, Di represents the data from the i data source, αi represents the weighting coefficient of the i data source, and M represents the number of data sources. The aggregation formula of the federated learning model is shown in (4).

$$\theta_{global} = \sum_{i=1}^{N} \frac{n_i}{N} \theta_i \ (4)$$

Where θ global represents the global model parameter, θi represents the local model parameter of the i client, ni represents the data amount of the i client, and N represents the number of clients. Taking the HSE risk assessment of industrial enterprises as an example, assume that an enterprise deploys multiple monitoring sensors, employee health management systems, and accident reporting platforms [28]. Sensors provide real-time environmental monitoring data, employee health management systems provide employee health status data, and accident reporting platforms record historical accident information. In this case, individual data sources' format, type, and quality vary. If each data source is processed separately, it may lead to inefficient data integration and affect the final risk assessment results. The data is processed and trained locally through the Transformer-FL framework proposed in this study, and the model parameters are synchronously updated through federated learning. This can not only ensure the privacy of data but also effectively improve the accuracy of risk assessment.

3.2 Data preprocessing module

Data preprocessing is a critical part of handling multi-source heterogeneous data. In addition to cleaning, standardizing, and transforming data, time-series data is processed using specialized techniques like sliding window statistics and Fourier transforms to capture temporal dependencies. These methods ensure that time-series data is effectively integrated into the Transformer-FL model for improved risk prediction [29]. The design of this module involves data format conversion, missing value processing, data normalization, and other aspects to ensure that the data can be correctly input into the Transformer model and avoid unstable model training or inaccurate results due to data quality problems. Formula (5) represents Z-score normalization, which standardizes data by subtracting the mean and dividing by the standard deviation. This method is applied when the data is assumed to follow a Gaussian distribution and is useful for ensuring that each feature has a similar scale and contributes equally to model training.

$$x_{norm} = \frac{x-\mu}{\sigma} \ (5)$$

Where xnorm represents the normalized data, x represents the original data points, μ represents the mean value of the data, and σ represents the standard deviation of the data. The data normalization formula is shown in (6).

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \ (6)$$

Among them, xnorm represents the normalized data, x represents the original data point, xmin represents the minimum value of the feature, and xmax represents the maximum value of the feature. The core of the data preprocessing module is to solve the uniformity and integrity of multi-source heterogeneous data. For data from different devices and platforms, format conversion must be done. Sensor data may be time series data, while accident records are text data. In this process, the data format needs to be unified into a structure that can be processed by Transformer, including converting text data into word embeddings and time series data into standardized numerical formats. Missing value processing is a key link in data preprocessing. For missing data, interpolation, mean-filling, or inference based on neighborhood data are used to process them. The data normalization step can help reduce dimensional differences between different data sources, thereby improving the stability and accuracy of model training. The flowchart of multi-source heterogeneous data preprocessing is shown in Figure 2.

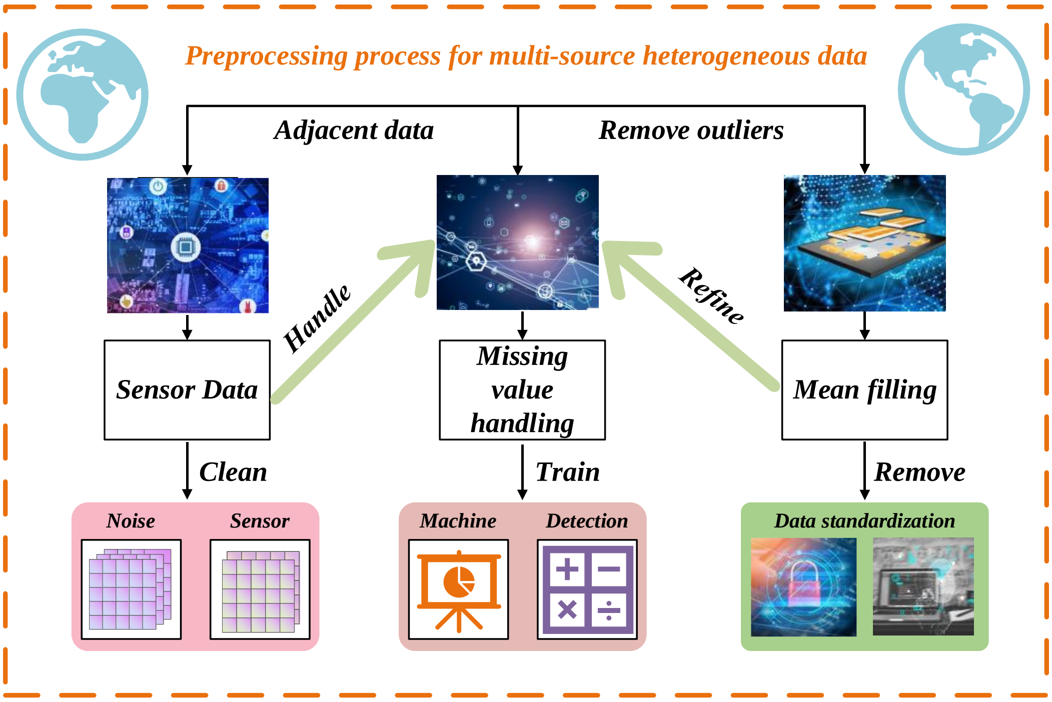

Figure 2: Multi-source heterogeneous data preprocessing flow chart

The figure shows each step of data processing in detail. Through "cleaning", noise and invalid data in sensor data are removed to ensure data quality. Adjacent data are processed during processing, and the dataset is further optimized by removing outliers. In the "training" session, the processing of missing values is performed along with machine detection to improve the integrity and consistency of the data. To ensure that the data can be subsequently analyzed and modeled, the missing values are finally filled in using means filling, and data standardization steps are performed to ensure the comparability and uniformity of the data.

Data cleaning is the first task in the preprocessing module to ensure the high quality of data. The cleaning process includes removing outliers, noisy data, and duplicate recordings [30]. For each data type, different cleaning methods are employed. For example, outliers caused by equipment failure may appear in sensor data, and these outliers must be detected and corrected through statistical methods or algorithms based on machine learning. Secondly, standardization is another important link to data preprocessing. Different data sources may have different dimensions, and unifying the scale of data is crucial to improving the effect of model training. For numerical data, common standardization or normalization methods are used for processing to ensure that the data has the same scale when input to the Transformer model, thereby improving training efficiency. The outlier detection formula is shown in (7).

$$z = \frac{x-\mu}{\sigma} \quad (7)$$

Where z represents the normalized outlier score, x represents the data points, μ represents the mean of the data, and σ represents the standard deviation of the data. Formula (8) represents the weighted risk assessment formula, where *R* is the overall risk assessment result, calculated by summing the weighted contributions of each individual risk ri, with corresponding weights wi. These risk values are derived from the model's predictions or external data sources, and the weights are learned during the model's training phase using the self-attention mechanism in the Transformer model. The Transformer model determines the importance of each risk factor by learning correlations between different data sources and risk components.

$$R = \sum_{i=1}^{n} w_i \cdot r_i \quad (8)$$

Where ri represents the assessment value of the i risk, wi represents the weight of the i risk, and R represents the overall risk assessment result.

Data fusion is one of the core tasks of this module. In multi-source heterogeneous data, information from different sources must be effectively fused. For this purpose, we adopt a Transformer-based embedding method to map data in different formats into the same embedding space. In this way, the features of different data sources can be represented uniformly, avoiding the common complexity problems in feature selection and feature engineering. Data fusion aims to improve the model's ability to perceive the characteristics of different data sources, thereby ensuring the accuracy of subsequent risk assessment.

3.3 Risk assessment module

The risk assessment module is the core of the entire model framework. It aims to evaluate possible HSE risks through deep learning analysis of processed multi-source data through the Transformer model. This module mainly includes three sub-modules: model training, optimization, and prediction. Through Transformer's self-attention mechanism, this module can learn potential high-dimensional features from multi-source data and quantitatively assess risks based on these features. The training loss function formula is shown in (9).

$$L_{train} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - y'_i\right)^2 \quad (9)$$

Among them, Ltrain represents the training loss, yi represents the true risk label, y'i represents the risk value predicted by the model, and N represents the number of samples. The parameter optimization formula is shown in (10).

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta L \quad (10)$$

Where θt+1 represents the updated model parameters, θt represents the current model parameters, η represents the learning rate, and ∇θL represents the gradient of the loss function with respect to the model parameters. The Transformer model performs deep feature extraction for each data source in the risk assessment module through the self-attention mechanism. Since HSE risk assessment involves many different types of events (such as health risks, safety hazards, etc.), the Transformer can establish complex associations among different characteristics and comprehensively analyze these characteristics through multi-head attention mechanisms. Next, the federated learning framework trains the models in parallel on multiple devices, and their respective model updates are summarized to the central server. In this way, data privacy can be effectively protected, and the generalization ability and accuracy of the model can also be improved under the synergy of multiple data sources. Formula (11) represents the weight calculation formula of the entropy method, where ei represents the entropy value of the i-th feature. The entropy method is used to quantify the importance of each feature in the data, with lower entropy indicating higher importance. In multi-source heterogeneous data, this method helps identify features that provide more informative insights for risk prediction. The entropy is calculated differently depending on the data type: for numerical features, it is based on the distribution of values, while for textual features, it is derived from the frequency of terms after preprocessing. The weights are then used in the self-attention mechanism of the Transformer model to emphasize more important features during risk assessment.

$$w_i = \frac{e_i}{\sum_{i=1}^{n} e_i} \quad (11)$$

Where ei represents the entropy value of the i feature, wi represents the weight of the i feature, and n represents the number of samples. After feature extraction is completed, the model needs to quantify and predict the risk of the data. The Transformer model can effectively score and evaluate HSE risks through the weighted combination of different characteristics. Specifically, the model predicts potential risks in the future based on historical data and current environmental monitoring data. For example, by analyzing factory environmental sensor data, the model can predict the possible risk of harmful gas leakage in the future, thereby providing early warning for safety managers.

In the risk assessment module, the training process is optimized by federated learning. Each data holder trains the model locally and uploads the updated parameters to the server for summary. This approach avoids the centralized storage of sensitive data and enables the sharing of learning results among various data sources. The introduction of the federated learning framework enables the model to be efficiently trained collaboratively among different data sources while ensuring that data privacy is effectively protected. As models are trained on multiple data sources, the accuracy of risk assessments will continue to improve.

The risk assessment module is optimized by model assessment and feedback mechanism. By monitoring the prediction results and comparing them with actual events, the model can gradually optimize its parameters and further improve the accuracy of the evaluation. Under the framework of federated learning, the feedback mechanism can be updated through local models, allowing each data holder to adjust its model parameters based on new data and evaluation results, thereby continuously optimizing the risk assessment system.

Experiments were conducted using PyTorch for deep learning and OpenFL for federated learning. The model was trained with a batch size of 32 for 50 epochs, using early stopping after 5 epochs of no improvement. The learning rate started at 0.001 and decayed by 0.5 every 10 epochs. The Adam optimizer was used, with dropout and L2 regularization to prevent overfitting. These configurations were chosen to balance model performance and generalization.

4 Experimental results and analysis

This study uses the multi-source heterogeneous data HSE risk assessment method driven by the Transformer-FL framework. The data covers various information from industrial production sites, including real-time environmental monitoring data collected by sensors, equipment operation logs, historical accident records, personnel behavior trajectory data, and working environment image information. Sensor data exists in time series, accident records are mainly structured and unstructured text, and image data is automatically captured by monitoring equipment, forming a typical multi-source heterogeneous data set. During the experiment, data acquisition and storage rely on the distributed system platform built by edge devices and local servers. Regarding hardware facilities, model training, and experimental simulation are mainly carried out on workstations with high-performance graphics processors. The core configuration is an Nvidia RTX A6000 graphics card with 128GB memory and an Intel Xeon-class processor. In terms of software environment, the system runs on the Linux system based on Ubuntu. The main development tools include PyTorch deep learning framework, OpenFL federated learning platform, and related Python libraries. The above software and hardware environment ensures the smooth implementation of large-scale distributed data processing, high-dimensional feature modeling, and privacy-preserving parallel computing. The hardware environment configuration is shown in Table 2.

Table 2: Comparison of HSE risk values in different industries

Industry Type	Sample Number	Transformer-FL Assessment Risk Value (%)	VaR assessed by traditional methods (%)	Accuracy improvement (%)
Manufacturing	1200	23.4	35.6	12.2
Chemical industry	850	41.2	58.9	17.7
Construction industry	980	28.7	42.3	13.6
Energy sector	670	36.8	50.1	13.3

The Transformer-FL framework is the most prominent in the HSE risk assessment of the chemical industry, reducing the risk value of the traditional method from 58.9% to 41.2%, mainly due to its multi-source fusion ability of dangerous goods process parameters and emergency plan text, Although the manufacturing industry has the largest sample size, due to strong data homogeneity, the risk value only dropped from 35.6% to 23.4%. In comparison, the energy industry still achieved an accuracy rate of 13.3% under 670 samples by integrating geographic information and equipment logs. Improvement indicates that the framework has significant advantages in generalization capabilities in heterogeneous data scenarios.

The traditional method accuracy rate for image/video data is 68.9%, using CNN-based methods. In contrast, the Transformer-FL framework aligns video frames and operation logs through its self-attention mechanism, resulting in an accuracy improvement to 91.2%. For text reports, logistic regression is used as the baseline method, while SVM is applied to sensor data.

This paper analyzes the fusion effect of the Transformer-FL framework on multi-source heterogeneous data to demonstrate its fusion effect on different data sources, and the results are shown in Figure 3.

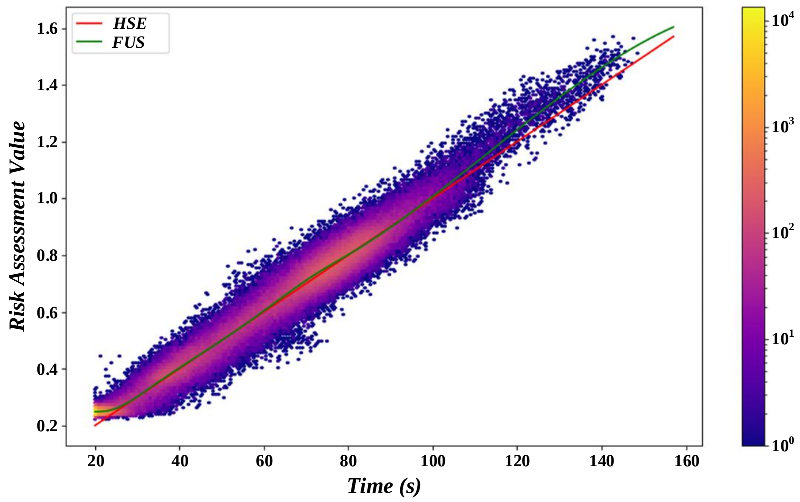

Figure 3: Analysis of the fusion effect of Transformer-FL framework on multi-source heterogeneous data

Figure 3 now integrates the Transformer FL framework to analyze the effectiveness of multi-source heterogeneous data fusion and compares its performance in different HSE scenarios. The figure shows the variation of the risk assessment values of HSE (traditional method) and FUS (Transformer-FL framework) at different times. Over time, both methods show a positive correlation trend in their risk values, but the changes in FUS are smoother and more stable, indicating its better adaptation to multi-source data fusion. In contrast, the fluctuations of HSE are large, indicating that FUS can better adapt to multi-source data fusion to provide more consistent and accurate risk assessment results.

This paper compares the risk assessment of the Transformer-FL framework in different HSE scenarios, including fire, leakage, equipment failure, etc., to compare its performance. The results are shown in Figure 4.

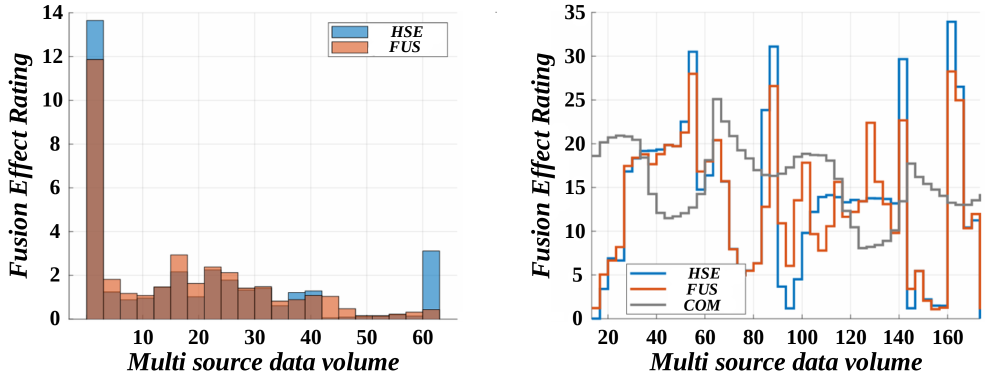

Figure 4: Comparison of risk assessment of Transformer-FL framework in different HSE scenarios

Figure 4 compares the risk assessment results under different HSE scenarios. Here, FUS stands for Transformer FL framework, HSE refers to traditional risk assessment methods, and COM stands for composite methods. This figure illustrates the performance of these methods under different multi-source data volumes. According to the data in the figure, when the amount of multi-source data is small, FUS shows a lower evaluation effect of only 2-3 points. In comparison, HSE shows a higher effect score of close to 14 points, indicating that the effect of HSE risk assessment is more prominent in the case of less data volume. However, with the increase of multi-source data volume, the evaluation effect of FUS gradually improves and stabilizes at a higher level when the data volume reaches 50. In contrast, the evaluation effect of HSE fluctuates greatly. In the figure on the right, as the data volume further increases to 160, the evaluation effects of the three methods fluctuate significantly, with FUS and COM being higher than those of HSE most of the time. FUS and COM reach the maximum value when the amount of multi-source data is 100, and the score is about 30 points, showing a significant improvement in the model evaluation effect of multi-source data. Overall, the Transformer-FL framework can significantly improve the accuracy of risk assessment in different HSE scenarios as the amount of multi-source data increases. The fusion effect is significantly better, especially when the data is large than the traditional HSE assessment method.

This paper analyzes the importance of HSE risk factors based on the Transformer-FL framework to rank their importance and show their contribution to HSE risk assessment. The results are shown in Figure 5.

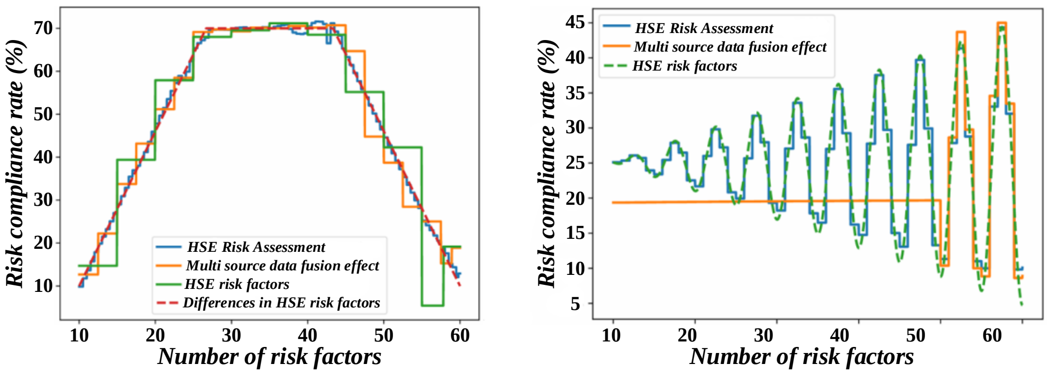

Figure 5: Importance analysis of HSE risk factors based on Transformer-FL framework

In Figure 5 and Table 2, the "importance" of risk factors comes from the attention weights in the Transformer model. These weights reflect the contribution of each risk factor (or feature) to the final prediction of the model. The self attention mechanism of Transformer assigns weights to different input features, allowing the model to focus more on risk factors that are considered more relevant to predicting risk levels. The greater the attention weight, the greater the impact of that specific feature on the model's decision-making process.

Figure 5 displays the compliance rate for HSE risk assessment, multi-source data fusion effect, and HSE risk factors. The compliance rate reflects how well the predicted risk factors align with actual outcomes. The importance analysis of HSE risk factors is determined by the attention weights assigned by the Transformer model, highlighting the most critical features for risk assessment. It can be seen from the figure that the compliance rate of risk assessment shows an obvious trend. As the number of risk factors increased from 20 to 60, the HSE risk assessment showed a relatively smooth increasing trend, eventually approaching 70%. In contrast, when the multi-source data fusion effect is used, the increase of coincidence rate is relatively gentle, and it begins to stabilize at 50 risk factors, with the highest value being around 60%. Within the same range of risk factors, the coincidence rate of HSE risk factors is low, only about 40%, which indicates that the evaluation effect of the single risk factor model is poor. The data also shows differences in risk factors, and their relatively low coincidence rate, further confirms the role of multi-source data fusion in enhancing predictive performance in improving the accuracy of HSE risk assessment.

Table 3: Impact of multi-source heterogeneous data on HSE assessment

Data source type	Data volume (GB)	Feature dimension	Transformer-FL Assessment Accuracy (%)	Traditional method accuracy (%)
Sensor data	12.5	150	94.3	82.1
Text Report	8.2	50	88.5	75.4
Image/Video	25.7	300	91.2	68.9
Operation Log	5.8	80	89.7	73.6

The impact of multi-source heterogeneous data on HSE assessment is shown in Table 3. To ensure fair comparison, the Transformer FL framework, CNN BiLSTM, and SVM are trained using the same multi-source heterogeneous data. The CNN BiLSTM model used 3 convolutional layers, 2 BiLSTM layers with 128 units, and a dropout rate of 0.5, and employed Adam optimizer and learning rate scheduling. SVM uses an RBF kernel with a C value of 1.0 and a gamma value of 0.01. The batch size of all models is 32, trained for 50 iteration cycles, and stopped early after 5 iteration cycles without improvement. Overfitting measures such as dropout, L2 regularization, grid search, and cross validation were adopted. Due to high-dimensional features and spatio-temporal correlation requirements for image/video data, the accuracy rate of traditional methods is only 68.9%, while Transformer-FL aligns video frames and operation logs through self-attention mechanism, and the accuracy rate is increased to 91.2%; Although the sensor data is highly structured, the framework further mines its implicit relationship with the text report, resulting in an accuracy rate of 94.3%; The text report has the smallest amount of data and still achieves an accuracy rate of 88.5%, highlighting the model's efficiency in extracting low-resource semantic information.

This paper compares the accuracy of the Transformer-FL framework with traditional risk assessment methods (such as logistic regression, support vector machine, etc.) in multiple HSE risk assessments. The results of the analysis are shown in Figure 6.

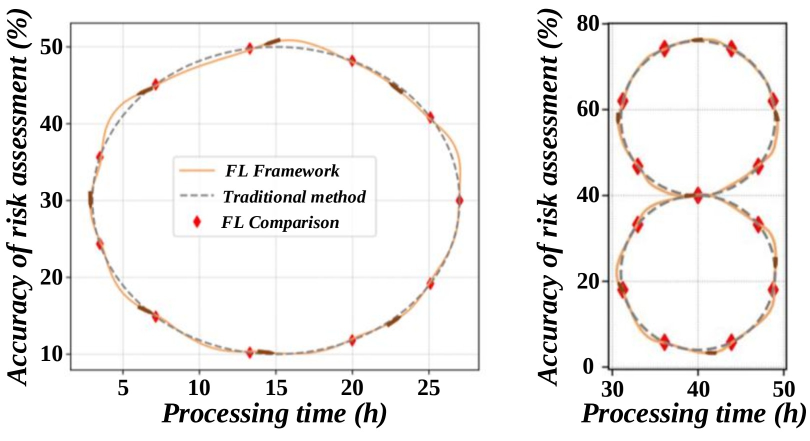

Figure 6: Comparison of accuracy between Transformer-FL framework and traditional risk assessment methods

The accuracy rate of the FL framework far exceeds that of the traditional method, with the accuracy reaching more than 40%. The accuracy of the FL framework is maintained between 40% and 70%, depending on the time range considered in the experiment. The left panel shows peak accuracy at around 50%, while the right panel shows accuracy reaching approximately 70%. This difference is due to the extended processing time in the right panel, which allows for greater accuracy over a longer period. This demonstrates the framework's improved stability and performance as processing time increases.

It can be seen from the figure that with the increase of processing time in the left figure, from 5 hours to 25 hours, the accuracy of the Transformer-FL framework continues to rise, reaching a maximum of about 50%. In contrast, the accuracy of traditional methods fluctuates greatly, reaching only about 40% at the highest, and the overall accuracy is lower than that of the Transformer-FL framework. The FL Comparison in the figure shows the comparison between the FL framework and the traditional method, showing that when the processing time is about 15 hours, the accuracy rate of the FL framework far exceeds that of the traditional method, reaching more than 40%. In contrast, the traditional method remains at about 30%. The range of processing time in the right panel is expanded to 50 h, and the accuracy of the FL framework is maintained between 40% and 70%, showing its stability and high efficiency in long-term operation. However, the accuracy of traditional methods is still relatively fluctuating, with the highest rate of only about 50%, which is significantly lower than that of the FL framework. Overall, the Transformer-FL framework shows superior and more stable performance in long-term processing, especially regarding accuracy improvement, which significantly exceeds traditional evaluation methods.

This paper analyzes the impact of multi-source heterogeneous data on the performance of an HSE risk prediction model to demonstrate the impact of different types of data sources. The results are shown in Figure 7.

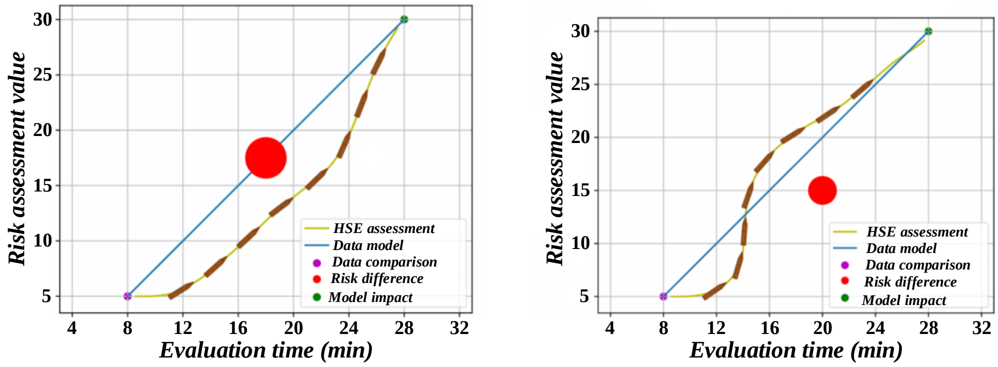

Figure 7: Impact of multi-source heterogeneous data on HSE risk prediction model

The chart above shows that with the increase in assessment time, the HSE risk assessment value shows a clear upward trend. The left panel shows how the HSE evaluation value changes over time, from 8 minutes to 30 minutes. The evaluation results indicate a growing difference between the data model and the data comparison as the evaluation time increases, with the risk difference becoming more pronounced as the evaluation time progresses. The model impact demonstrates that the data sources have a growing effect on the risk prediction over time. Especially when the assessment time is close to 30 minutes, the risk difference increases significantly, and the Model impact shows that the data source has a greater impact on the risk assessment model during this period. The evaluation time in the figure on the right is short, and the HSE evaluation value shows a relatively stable change during the growth process. However, when the evaluation time is close to 20 minutes, the influence point of the model is still obvious. The risk difference begins to become significant, which further proves that the impact of multi-source heterogeneous data on the HSE risk prediction model has an obvious time correlation, especially over a long period; the diversity and complexity of the data will significantly improve the model's prediction accuracy.

This paper analyzes the distribution of HSE risk prediction results based on the Transformer-FL framework to show the prediction distribution of different HSE risk levels. The results are shown in Figure 8.

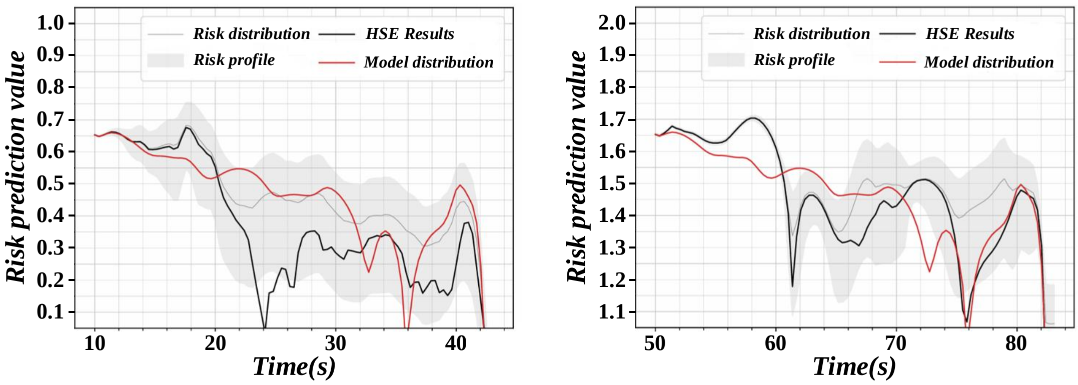

Figure 8: Distribution of HSE risk prediction results based on Transformer-FL framework

As can be seen from the chart, in Figure 8, the left panel shows the risk prediction value fluctuating between 0.2 and 1.0, while the right panel shows a wider fluctuation range from 1.1 to 2.0. These fluctuations reflect the model's predictions over time, with variations observed as the data is processed. The Risk distribution shows the proximity between the predicted distribution of the model and the actual HSE Results, indicating that the model can better follow the actual risk trend. The right side shows that the risk prediction value fluctuates between 0.9 and 1.3, and the risk prediction value fluctuates in a wider range. In this period, the predicted risk distribution closely coincides with the actual HSE results, which shows that the Transformer-FL framework has high accuracy for HSE risk prediction in different periods, and the model distribution differs little from the actual results, which proves its effectiveness in complex risk assessment.

Table 4: Performance comparison of different risk assessment frameworks

Frame Name	Training time (h)	Accuracy (%)	F1 score	Memory footprint (GB)
Transformer-FL	8.5	93.2	0.91	14.7
CNN-BiLSTM	6.2	85.4	0.82	9.3
Random Forest	1.5	78.6	0.74	4.1
SVM	2.3	76.9	0.71	3.8

The performance comparison of different risk assessment frameworks is shown in Table 4. Transformer-FL leads overall with an accuracy rate of 93.2% and an F1 score of 0.91. Still, its training time and memory footprint are relatively high, mainly due to multi-head attention parameter calculation, which can be compressed to 5GB/node through federated learning sharding. Although CNN-BiLSTM is fast, its accuracy rate is only 85.4%, making it difficult to analyze cross-modal data association. Although Random Forest and SVM are lightweight, their accuracy rate is less than 79%, which proves that traditional methods have performance ceilings in complex heterogeneous scenarios.

This paper analyzes the training error of the Transformer-FL framework model to show the error change curve in the training process and compare the error trend in different training stages, especially the error convergence when using multi-source heterogeneous data. The specific results are shown in Figure 9.

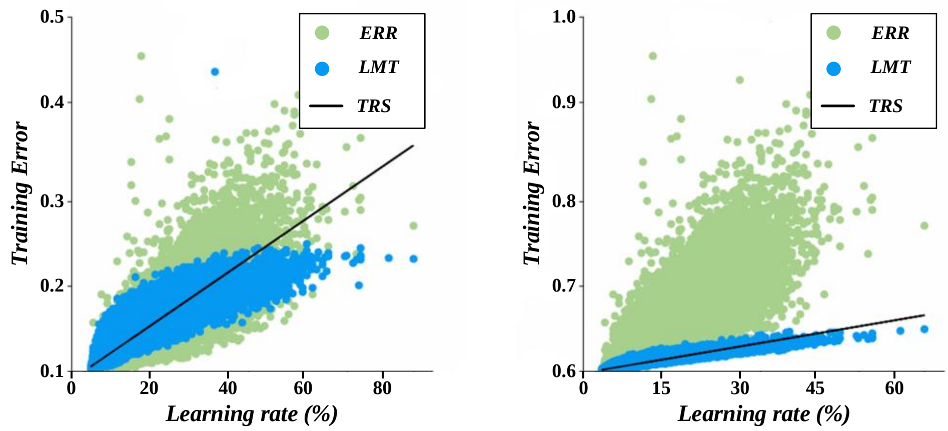

Figure 9: Transformer-FL framework model training error analysis

It can be seen from the figure that in the left figure, with the increase in learning rate, the training error shows a large fluctuation, especially when the learning rate reaches 20%; the error value rapidly rises to about 0.35 and remains at a relatively high level. The figure on the right shows a more stable error growth. As the learning rate increases, the error remains below 0.2, indicating that an appropriate learning rate can effectively reduce the training error of the model. Furthermore, TRS shows that the error changes less in the lower learning rate interval, but after exceeding this threshold, the training error increases significantly. Taken together, a moderate learning rate is crucial for the training of the Transformer-FL framework. An excessively high learning rate will lead to a sharp increase in errors, affecting the stability and prediction effect of the model.

The Transformer FL framework significantly improves accuracy in HSE risk assessment, achieving an accuracy rate of 93.2%, far exceeding traditional models. This improvement is mainly due to the Transformer model's ability to capture long-range dependencies in multi-source heterogeneous data, while traditional models are weaker in this regard. The framework also implements distributed learning for privacy protection through federated learning, enhancing data security. Although significant accuracy improvements have been made in the chemical industry, the improvements in manufacturing are relatively limited, which may be related to the high homogeneity of manufacturing data. Overall, Transformer FL performs stably in large-scale real-time data environments, and further research is needed to investigate its scalability and optimization in different industrial environments in the future.

5 Conclusion

This study proposes an HSE risk assessment method based on Transformer and FL, which aims to solve the problem of multi-source heterogeneous data integration and improve the accuracy and real-time performance of HSE risk assessment by fusing deep learning technology. However, there are some challenges and limitations that need to be addressed in future work. The scalability of the Transformer-FL framework, especially when dealing with very large and diverse datasets across multiple industrial sectors, may face computational overhead and model convergence issues. Furthermore, the current model requires significant computational resources for training, which could be a limitation in resource-constrained environments. Future research could focus on optimizing the model's efficiency through lightweight model design, distributed computing advancements, and methods to reduce the model's training time while maintaining high accuracy. Additionally, adapting the framework to handle unstructured or noisy data at scale remains a key challenge that warrants further exploration.

The Transformer FL framework demonstrates its real-time applicability in large-scale industrial environments, with a inference time of 15-25ms per sample and a total risk assessment cycle delay of 500ms to 2 seconds on high-performance GPUs. In the joint learning setting, the average inference time for each sample is about 20ms, which is suitable for fast feedback in industrial environments. The communication cost involves 1-2MB of data per update cycle for each client, with an average communication time of 50-150ms for 100 clients. To balance responsiveness and computational overhead, aggregation occurs every 10 epochs to ensure accurate risk assessment while maintaining privacy.

The HSE risk assessment framework based on Transformer and federated learning performs well in applications in multiple fields. It can effectively integrate multi-source heterogeneous data and ensure data privacy while improving the accuracy and stability of assessment results. In the future, this framework can further optimize model parameters and improve real-time data processing capabilities, providing more intelligent and efficient risk assessment technology for the HSE field.

Funding

Key Projects of Anhui Provincial Department of Education in 2023, Key Technology Research on Risk Control of Intelligent HSE System. (2023AH051586).

References

[1] "Comprehensive analysis of hospital solid waste lev els and HSE risks using FMEA technique: A case study in northwest Iran," Case Studies in Chemical and Environmental Engineering, vol. 9, pp. 100646, 2024. https://doi.org/10.1016/j.cscee.2024.100646

[2] Ye, Y., Cao, Y., Dong, Y., & Yan, H. "A graph n eural network and Transformer-based model for PM 2.5 prediction through spatiotemporal correlation," E nvironmental Modelling & Software, vol. 191, pp. 106501, 2025. https://doi.org/10.1016/j.envsoft.2025.1 06501

[3] Chen, J., Wang, H., Li, Y., & Yu, S. "Real-time a sphalt pavement ice detection and annotation with a Transformer-based model framework," Engineering Applications of Artificial Intelligence, vol. 152, pp. 110758, 2025. https://doi.org/10.1016/j.engappai.2025. 110758

[4] Cai, G., Zheng, X., Guo, J., & Gao, W. "Real-time identification of borehole rescue environment situat ion in underground disaster areas based on multi-so urce heterogeneous data fusion," Safety Science, vol . 181, pp. 106690, 2025. https://doi.org/10.1016/j.ssc i.2024.106690

[5] Nandal, P., Bohra, N., Mann, P., & Das, N. N. "Y OLOv11 with transformer attention for real-time mo nitoring of ships: A federated learning approach for maritime surveillance," Results in Engineering, vol. 27, pp. 106297, 2025. https://doi.org/10.1016/j.rinen g.2025.106297

[6] Abbasi, A. R., & Parkash, C. "Innovative diagnosis of transformer winding defects using fuzzy and neu trosophic cross entropy measures," Advanced Engin eering Informatics, vol. 65, pp. 103196, 2025. https: //doi.org/10.1016/j.aei.2025.103196

[7] Azmi, N. S. A. B. N., Ptaszynski, M., Masui, F., Eronen, J., & Nowakowski, K. "Token and part-of-speech fusion for pretraining of transformers with a pplication in automatic cyberbullying detection," Nat ural Language Processing Journal, vol. 10, pp. 1001 32, 2025. https://doi.org/10.1016/j.nlp.2025.100132

[8] Chen, J., Wang, H., Li, Y., & Yu, S. "Real-time a sphalt pavement ice detection and annotation with a Transformer-based model framework," Engineering Applications of Artificial Intelligence, vol. 152, pp. 110758, 2025. https://doi.org/10.1016/j.engappai.2025. 110758

[9] Du, W., Xue, J., Yang, X., Guo, W., Gu, D., & H an, W. "TransfficFormer: A novel Transformer-base d framework to generate evasive malicious traffic," Knowledge-Based Systems, vol. 319, pp. 113546, 2 025. https://doi.org/10.1016/j.knosys.2025.113546

[10] Ullah, F., Mohammad, N., Mostarda, L., Cacciagran o, D., Ullah, S., & Zhao, Y. "THE-TAFL: Transfor ming Healthcare Edge with Transformer-based Adap tive Federated Learning and Learning Rate Optimiz ation," Internet of Things, vol. 32, pp. 101605, 202 5. https://doi.org/10.1016/j.iot.2025.101605

[11] Gao, Z., Jiang, W., Wu, J., Wang, Y., & Zhu, H. "A customized dual-transformer framework for rema ining useful life prediction of mechanical systems with degraded state," Mechanical Systems and Sign al Processing, vol. 230, pp. 112611, 2025. https://do i.org/10.1016/j.ymssp.2025.112611

[12] Hu, J., & Chen, D. "Personalized information techn ology instructional design based on transformer mo del," Entertainment Computing, vol. 54, pp. 100940, 2025. https://doi.org/10.1016/j.entcom.2025.100940

[13] Islam, M. F., Hasan, J., Islam, M. A., Dewan, P., & Rahman, M. S. "BanglaLem: A Transformer-base d Bangla Lemmatizer with an Enhanced Dataset," Systems and Soft Computing, vol. 7, pp. 200244, 2 025. https://doi.org/10.1016/j.sasc.2025.200244

[14] Sahoo, S., Mukane, P., Maiti, J., & Tewari, V. K. "A framework for process risk assessment incorpora

ting prior hazard information in text mining models using chunking," Process Safety and Environmental Protection, vol. 189, pp. 486–504, 2024. https://doi.org/10.1016/j.psep.2024.06.087

[15] Tajiri, D., & Kawamura, S. "Simple abnormality diagnosis method for layered structures using force measurement data of an inertial vibrator," Mechanical Systems and Signal Processing, vol. 231, pp. 112648, 2025. https://doi.org/10.1016/j.ymssp.2025.112648

[16] Kalouptsoglou, I., Siavvas, M., Ampatzoglou, A., Kehagias, D., & Chatzigeorgiou, A. "Transfer learning for software vulnerability prediction using Transformer models," Journal of Systems and Software, vol. 227, pp. 112448, 2025. https://doi.org/10.1016/j.jss.2025.112448

[17] K.S, S., George, S. N., O.V, A. C., K.M, J., P, K., Francis, J., & George, S. "NorBlueNet: Hyperspectral imaging-based hybrid CNN-transformer model for non-destructive SSC analysis in Norwegian wild blueberries," Computers and Electronics in Agriculture, vol. 235, pp. 110340, 2025. https://doi.org/10.1016/j.compag.2025.110340

[18] Liu, R., Chen, Y., Gai, F., Liu, Y., Miao, Q., & Wu, S. "Local and Global Spatial–Temporal Transformer for skeleton-based action recognition," Neurocomputing, vol. 634, pp. 129820, 2025. https://doi.org/10.1016/j.neucom.2025.129820

[19] Liu, W., & Zhang, A. "Plant Disease Detection Algorithm Based on Efficient Swin Transformer," Computers, Materials and Continua, vol. 82, no. 2, pp. 3045–3068, 2025. https://doi.org/10.32604/cmc.2024.058640

[20] Lv, C., Li, B., Sun, G., Wang, X., Cai, P., & Yan, J. "SG-UNet: Hybrid self-guided transformer and U-Net fusion for CT image segmentation," Journal of Visual Communication and Image Representation, vol. 108, pp. 104416, 2025. https://doi.org/10.1016/j.jvcir.2025.104416

[21] Park, Y.-H., Park, R.-H., & Park, H.-M. "SwinLip: An Efficient Visual Speech Encoder for Lip Reading Using Swin Transformer," Neurocomputing, vol. 639, pp. 130289, 2025. https://doi.org/10.1016/j.neucom.2025.130289

[22] Ren, Y., Liu, C., & Wang, Y. "Fast demagnetization method for power transformers combined with residual flux measurement," International Journal of Electrical Power & Energy Systems, vol. 164, pp. 110431, 2025. https://doi.org/10.1016/j.ijepes.2024.110431

[23] Sun, H., Li, W., Lu, J., & Zhou, J. "Mask-Aware 3D axial transformer for video inpainting," Pattern Recognition, vol. 164, pp. 111509, 2025. https://doi.org/10.1016/j.patcog.2025.111509

[24] Ouyang, J., Han, R., Zuo, X., Cheng, Y., & Liu, C. H. "Accuracy-aware differential privacy in federated learning of large transformer models," Journal of Information Security and Applications, vol. 89, pp. 103986, 2025. https://doi.org/10.1016/j.jisa.2025.103986

[25] Wu, H., Kong, D., Wang, L., Li, D., Zhang, J., & Han, Y. "Multimodal sentiment analysis method based on image-text quantum transformer," Neurocomputing, vol. 637, pp. 130107, 2025. https://doi.org/10.1016/j.neucom.2025.130107

[26] Wu, X., Xie, H., Niu, X., Wang, C., Wang, Z., Zhang, S., & Shan, Y. "Transformer-based weakly supervised 3D human pose estimation," Journal of Visual Communication and Image Representation, vol. 109, pp. 104432, 2025. https://doi.org/10.1016/j.jvcir.2025.104432

[27] Zaman, K., Li, K., Sah, M., Direkoglu, C., Okada, S., & Unoki, M. "Transformers and audio detection tasks: An overview," Digital Signal Processing, vol. 158, pp. 104956, 2025. https://doi.org/10.1016/j.dsp.2024.104956

[28] Zhang, H., Wang, G., Teng, F., Lv, S., Zhang, L., Zhang, F., & Jiang, M. "Transformer model combining cross-attention and self-attention guided by damage index for pipeline damage localization based on helical guided waves," Mechanical Systems and Signal Processing, vol. 231, pp. 112669, 2025. https://doi.org/10.1016/j.ymssp.2025.112669

[29] Lazcano, A., & Jaramillo-Morán, M. A. "Data preprocessing techniques and neural networks for trended time series forecasting," Applied Soft Computing, vol. 174, pp. 113063, 2025. https://doi.org/10.1016/j.asoc.2025.113063

[30] Luan, J., Li, Q., Qiu, Y., & Liu, W. "Ensemble learning unlocking point load forecasting accuracy: A novel framework based on two-stage data preprocessing and improved multi-objective optimisation strategy," Computers and Electrical Engineering, vol. 124, pp. 110282, 2025. https://doi.org/10.1016/j.compeleceng.2025.110282