Article

Not peer-reviewed version

Intelligent Tunnel Fire Detection Technology Based on the Large Language Model Multi-Agent Collaboration

Ding Zeng * , Ao Gao * , Zhisheng Xu

Posted Date: 10 April 2026

doi: 10.20944/preprints202604.0757.v1

Keywords: tunnel fire; computational fluid dynamics; large language models; multi-agent collaboration; CNN-LSTM

doi:10.20944/preprints202604.0757.v1

Article

Intelligent Tunnel Fire Detection Technology Based on the Large Language Model Multi-Agent Collaboration

Ding Zeng [1,*], Ao Gao [2,*] and Zhisheng Xu [1]

[1] Institute of Disaster Prevention Science and Safety Technology, Central South University, Changsha 410075, China

[2] Shenyang Fire Research Institute of MEM (National Engineering Research Center for Fire and Emergency Rescue, Shenyang 110034, China

* Correspondence: dingzeng@csu.edu.cn (D.Z.); devingao@126.com (A.G.)

Abstract

To address the issues of manual operation dependency and low efficiency in tunnel fire research combining computational fluid dynamics (CFD)with deep learning, this paper proposes a multi-agent collaborative framework based on large language models to automate the entire process of inverting fire source characteristics. The framework decomposes the traditional workflow into four specialized agents, namely physical modeling, data governance, model training, and evaluation analysis, which collaboratively execute end-to-end tasks from CFD scenario generation to model deployment. The results demonstrate that the CNN-LSTM model performs optimally. Under a 6 second observation window and 10 meter sensor spacing, the average R^2 reaches 0.942, representing a 2% improvement over the baseline LSTM model, while the RMSE is reduced by 28.8%. Under sparse deployment with 30 meter spacing, the average R^2 remains as high as 0.917, validating the effectiveness of integrating spatial feature extraction with temporal modeling. This work provides an efficient technological pathway for intelligent tunnel fire identification and advances the research paradigm from manual optimization to multi-agent system optimization.

Keywords: tunnel fire; computational fluid dynamics; large language models; multi-agent collaboration; CNN-LSTM

---

1. Introduction

Tunnels serve as critical infrastructure in modern urban and intercity transportation networks, making operational safety paramount. However, the inherently long, narrow, and enclosed nature of tunnel structures presents severe challenges in the event of a fire [1]. The rapid spread of high-temperature toxic smoke within the confined space not only poses a serious threat to personnel safety but also places significant demands on fire rescue decision-making [2,3]. Therefore, the ability to quickly and accurately identify the fire source's location, scale, and ventilation-affected state after ignition holds irreplaceable scientific value and engineering urgency for guiding evacuation route planning, optimizing firefighting resource deployment, and minimizing loss of life and property [4,5].

Conventional tunnel fire detection technologies, such as smoke detectors based on smoke concentration thresholds, video image systems reliant on clear visibility, and linear heat-sensitive cables monitoring temperature changes, commonly suffer from inherent limitations in complex real-world tunnel environments. (e.g., exhaust gas interference, low visibility smoke, fire plume deflection due to strong longitudinal airflow). These limitations include high false alarm rates, significant localization delays, and high maintenance costs [6]. In recent years, with advancements in computational power and artificial intelligence technologies, methods integrating Computational

Fluid Dynamics (CFD) numerical simulation with data-driven deep learning have opened new technical pathways for the intelligent inversion and prediction of tunnel fire characteristics [7].Such approaches construct CFD fire scenario databases covering multiple operating conditions and train deep learning models to learn the complex nonlinear mapping relationships between sensor time-series data and key fire parameters, thereby achieving inversion of fire source characteristics.

Early research widely applied the Back Propagation Neural Network (BPNN) for fire parameter identification. Xue et al. [8] first utilized BPNN to achieve tunnel fire location identification based on smoke concentration, temperature, and CO sensor data. Sun et al. [9] further validated the feasibility of BPNN for fire localization through full-scale experiments based on temperature data. Barros-Daza et al. [10] extended BPNN to mine fire scenarios, achieving high-precision classification of fire source locations based on 500 numerical simulation scenarios. These studies verified the feasibility of neural network modeling; however, BPNN flattens time-series data, making it difficult to capture evolutionary patterns along the temporal dimension [11].

With the advancement of deep learning, Convolutional Neural Networks (CNNs) have garnered attention for their powerful spatial feature extraction capabilities. Allaire et al. [12] applied CNNs to surrogate modeling of wildfire spread simulation, demonstrating the ability of deep learning to fit complex physical fields. Meanwhile, Long Short-Term Memory (LSTM) networks have provided an effective tool for processing fire time-series data [13]. Wu et al. [14] first constructed a large-scale numerical scenario database for tunnel fires based on FDS, utilizing LSTM to achieve continuous regression prediction of multiple parameters, including fire source location and heat release rate. They also conducted sensitivity analysis on sensor deployment strategies, indicating that prediction accuracy meets engineering requirements when sensor spacing does not exceed 20 meters. In tunnel scenarios, Zhang et al. [5] systematically compared the performance of BPNN, CNN, LSTM, and CNN-LSTM in tasks involving fire source localization and different damper combinations, finding that hybrid models integrating spatial and temporal features have a clear advantage in classification tasks. However, their study did not investigate continuous regression prediction for tunnel fires.

Although the artificial intelligence-driven research paradigm for tunnel fires has achieved significant progress, it still faces critical bottlenecks [15]. First, the low degree of automation in the research workflow leads to high labor and time costs. Constructing physically credible CFD models requires deep domain expertise, and the research involves substantial repetitive manual tasks such as geometric modeling, meshing, and solver configuration, which are cumbersome and error-prone [16]. Second, the management of sensor time-series data is complex, and preprocessing is labor-intensive. The process of extracting, cleaning, and organizing data into structured training sets is time-consuming and technically demanding. Finally, the lengthy experimental research cycle makes it difficult to conduct large-scale, systematic parametric sensitivity analysis and iterative optimization of model architectures. This “artisanal” research model has become a core obstacle restricting the rapid development and engineering application of the field [17] .

To address the aforementioned bottlenecks, this paper proposes and constructs a multi-agent collaborative artificial intelligence framework for fire source characteristic inversion in tunnel fires [18]. This framework introduces an agent system based on Large Language Models (LLMs). By simulating the collaborative model of a scientific research team, this framework restructures the traditional workflow, which spans physical modeling, data production, model training, and comprehensive evaluation and previously relied heavily on manual operations. It thereby transforms this process into a highly automated and intelligent collaborative workflow. Its core objective is to restructure the existing research paradigm, freeing researchers from tedious repetitive tasks and enabling them to focus on higher-level scientific problem definition and innovative exploration [19,20].

The main contributions of this paper are summarized as follows:

(1) This study presents the first implementation that integrates an LLM-based multi-agent collaborative system with domain-specific knowledge of tunnel fires, achieving end-to-end automation of the fire source characteristic inversion process from high-level task description to final

doi:10.20944/preprints202604.0757.v1

model delivery. This encompasses core aspects such as tunnel physical modeling, data governance, model training, and evaluation feedback, thereby significantly enhancing research efficiency in tunnel fire studies.

(2) This study systematically compares and reveals the performance differences of various deep learning architectures in the task of fire source characteristic inversion. Based on a database of 100 tunnel fire scenarios simulated using FDS, this paper comprehensively evaluates the performance of five representative models in multi-task continuous regression for fire source location, heat release rate, and longitudinal ventilation velocity. By analyzing their capabilities in temporal feature extraction, spatial feature fusion, and long-range dependency capture, this study provides in-depth insights into the underlying mechanisms and applicable scenarios of different architectures, offering a reliable basis for model selection in future research.

(3) This study validates the superior performance and strong robustness of CNN-LSTM under the challenge of continuous multi-task prediction. Under a sparse sensor deployment condition with a spacing of 30 m, the CNN-LSTM model achieves a validation set R^2 of 91.7% under the early time window of 6 seconds, outperforming the baseline LSTM model by 1.6%.

2. Multi-Agent Collaborative Research Framework

The core design philosophy of the multi-agent collaborative research framework is to decouple the complex research workflow into a series of standardized and modular tasks. Those tasks executed autonomously by specialized and collaborative agents, thereby constructing an end-to-end automated pipeline from physical problem definition to AI model delivery.

Figure 1 illustrates the overall architecture and agent collaboration flow of the framework. The framework consists of four core agents, including Physical Modeling Agent, Data Governance Agent, Model Training Agent, and Evaluation Analysis Agent. Upon receiving a high-level task description from the user, the framework automatically executes the corresponding workflow, ultimately outputs trained models and a comprehensive evaluation report.

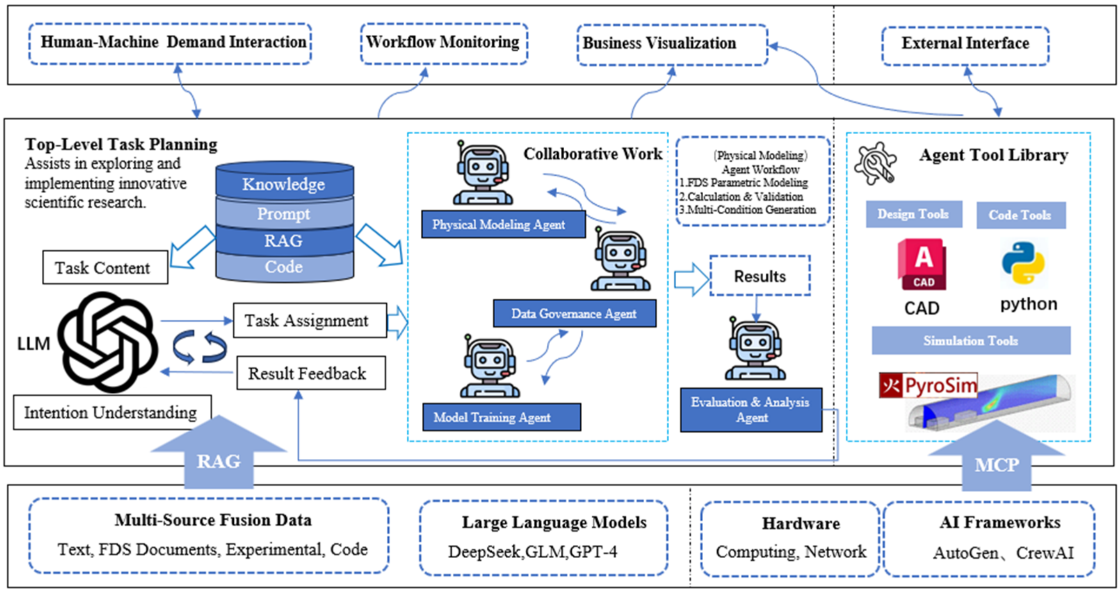


Figure 1. Multi-Agent Collaborative Research Framework for Tunnel Fire.

2.1. Physical Modeling Agent

The Physical Modeling Agent is responsible for converting human expert research requirements described in natural language into physically credible and computationally viable CFD model libraries. Based on Large Language Model mechanisms, it identifies, extracts, and formalizes key requirements from expert-provided descriptions. Figure 2 illustrates the workflow of the three agents involved in physical modeling.

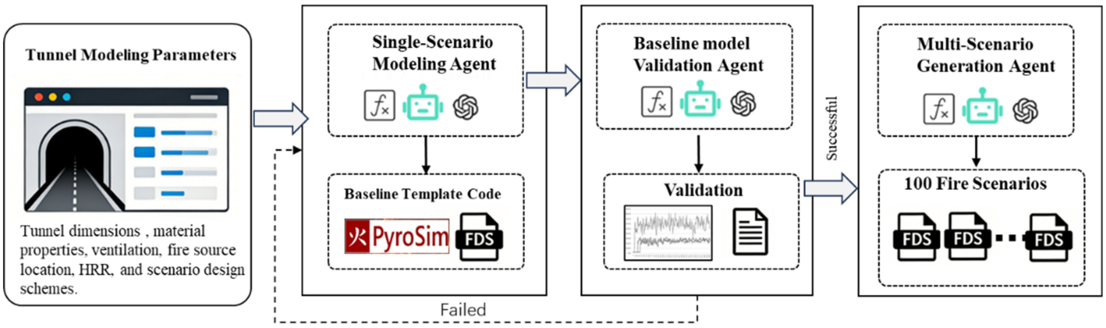
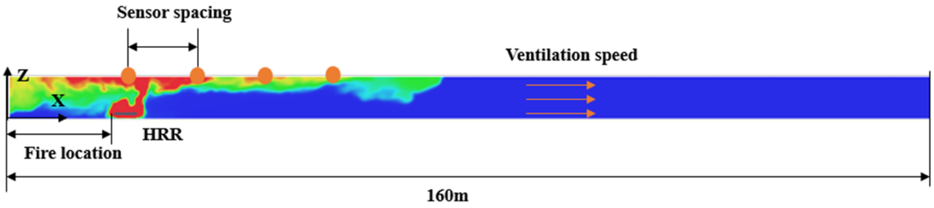
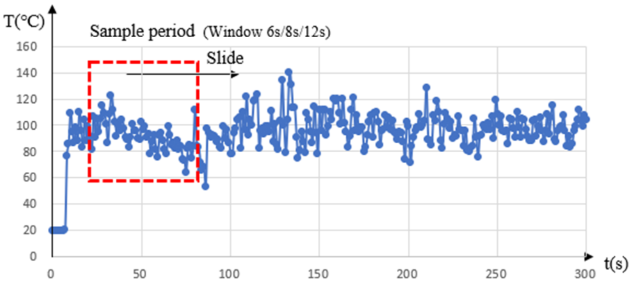


Figure 2. Schematic diagram of Tunnel Physical Modeling Agent.

Single-Case Modeling Agent acts as the designer for tunnel modeling, responsible for identifying and extracting user descriptions of tunnel geometry, materials, boundary conditions, and initial fire source parameters. Combined with constraints from a dedicated tunnel fire knowledge base, it automatically generates physically plausible FDS files for tunnel modeling [21].This study takes the full-scale road test tunnel of the Sichuan Fire Research Institute in China as an example [14]. The model tunnel is set with a length of 160 m (x-direction), a width of 6 m (y-direction), and height of 6 m (z-direction). Tunnel surfaces are set as inert walls, and the ambient temperature is fixed at 20 °C. The left exit is open, while the right exit is set as open or constant ventilation boundary depending on ventilation conditions. The inner tunnel surface is assumed to be perfectly smooth (roughness=0). The burning vehicle is modeled using a rectangular igniter (length 3 m × width 6 m, positioned 0.5 m above the ground) to simulate a vehicle fire.



(a) Key Parameters of Tunnel Fire Numerical Model



(b) Sampling Period Window

Figure 3. Single-Case Modeling Schematic (a) Key Parameters of Tunnel Fire Numerical Model (b) Sampling Period Window.

Benchmark Template Validation Agent is responsible for verifying the accuracy of single-case FDS simulation results. This agent can automatically execute the FDS file for a specified case, extract the temperature curves at left, middle, and right exits of the tunnel ceiling from the generated temperature field, and compare them with benchmark results from full-scale fire experiments or authoritative literature to validate model agreement [22]. Successfully validated FDS templates are passed as benchmark templates to the next stage, ensuring that the subsequent research is built upon a reliable physical foundation.

Multi-Case Generation Agent acts as the experimental design engine. Based on user-defined experimental design schemes, this agent automatically writes scripts to batch modify key parameters in input files, generating a complete library of FDS case files that cover all parameter combinations. As shown in Table 1, this study considers three key variables, including fire source location (X), heat release rate (Q), and longitudinal ventilation velocity (V). A total of 100 independent fire cases are generated, laying the foundation for subsequent large-scale parametric studies.

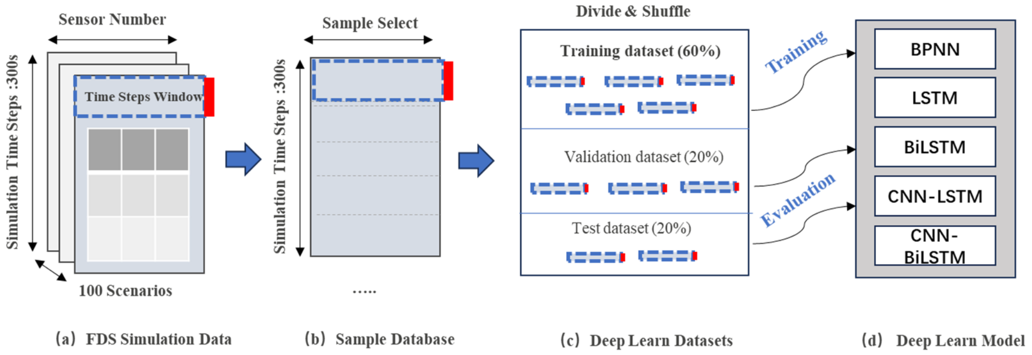
Table 1. Design of 100 Tunnel Fire Scenarios.

Core Variable	**Symbol**	**Number**	**Specific Gradient Values**
Fire Source Location	X	5	16m(0.1L)、32m(0.2L)、64m(0.4L)、96m(0.6L)、128m(0.8L)（L=160m）
Heat Release Rate（HRR）	Q	5	5MW、10MW、20MW、50MW、80MW
Longitudinal Ventilation Velocity	V	4	0m/s、1m/s、2m/s、4m/s

2.2. Data Governance Agent

The Data Governance Agent acts as a big data engineer, responsible for efficiently and reliably producing high-quality datasets suitable for deep learning. It interacts with high-performance computing clusters to intelligently package generated FDS job files, queue tasks for batch submission, and monitor job status and resource usage in real time. Upon computation completion, it automatically locates and parses the output files for each case, accurately extracting temperature time-series data from all preset measurement points to form a raw data pool.

Based on FDS numerical simulations, a dedicated dataset comprising 100 independent tunnel fire cases was constructed. For each case, one temperature measurement point is placed at equal longitudinal spacing along the tunnel ceiling, with spacing set to 10 m, 20 m, or 30 m. Temperature variations during the first 300 seconds after fire ignition are recorded at a sampling frequency of 1 Hz, forming temperature time-series. The dataset is randomly split into training, test, and validation sets in a ratio of 6:2:2 , ensuring balanced distribution of case parameters across the splits [23].



(a) FDS Simulation Data (b) Sample Database (c) Deep Learn Datasets (d) Deep Learn Model

Figure 4. Schematic Diagram of Tunnel Fire Scenario Simulation Database Design.

2.3. Model Training Agent

In fire parameter inversion research, traditional experimental workflows typically rely on manual execution of repetitive tasks such as data preprocessing, model selection, hyperparameter tuning, and result aggregation. Given the combination space involving multiple window lengths, multiple sensor spacings, and multiple models, such manual operations are not only inefficient but also make it difficult to ensure experimental fairness and reproducibility.

To address these issues, this paper designs and implements an AI agent tailored for multi-output continuous regression tasks. Through a modular design, this agent takes raw temperature time-series data as input and automatically executes data loading, data preprocessing, model construction, training evaluation, and result summarization. It systematically traverses all experimental configurations, providing a stable, efficient, and scalable infrastructure for fire parameter inversion research.

2.3.1. Model Selection

To systematically evaluate the performance differences of various deep learning architectures in fire parameter inversion tasks, this study selects five representative models, ranging from simple fully-connected networks to complex spatiotemporal hybrid architectures [24].

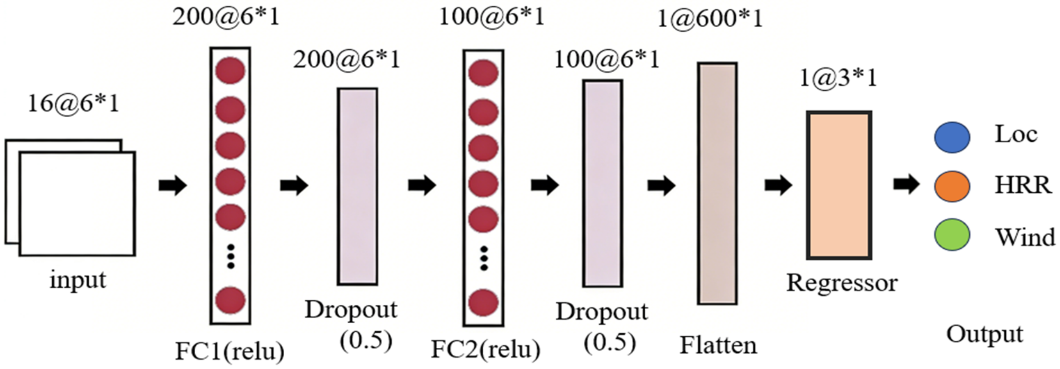
BPNN：Flattens spatiotemporal data and uses fully connected layers for direct regression, serving as a baseline model without awareness of temporal structure.

LSTM：Captures long-term and short-term dependencies in time series through gating mechanisms, serving as a classic benchmark for processing sequential data.

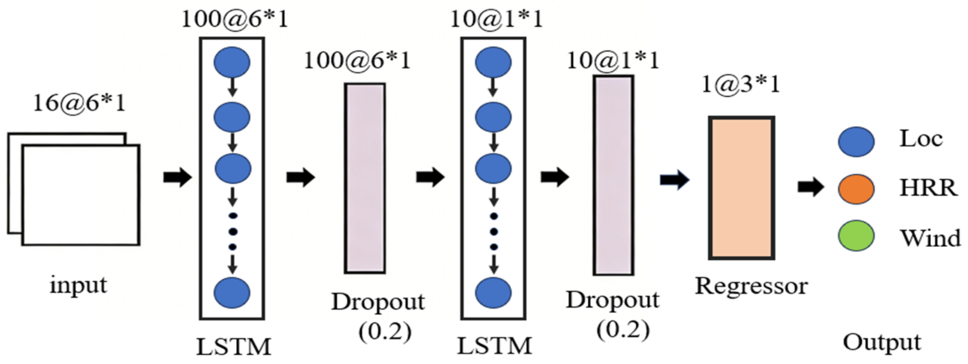
BiLSTM: Utilizes both forward and backward temporal information simultaneously to enhance comprehensive understanding of temporal context.

CNN-LSTM: Fully leverages the local perception capability of CNNs and the long-short-term memory characteristics of LSTMs, enhancing the model’s ability to understand and predict complex spatiotemporal data.

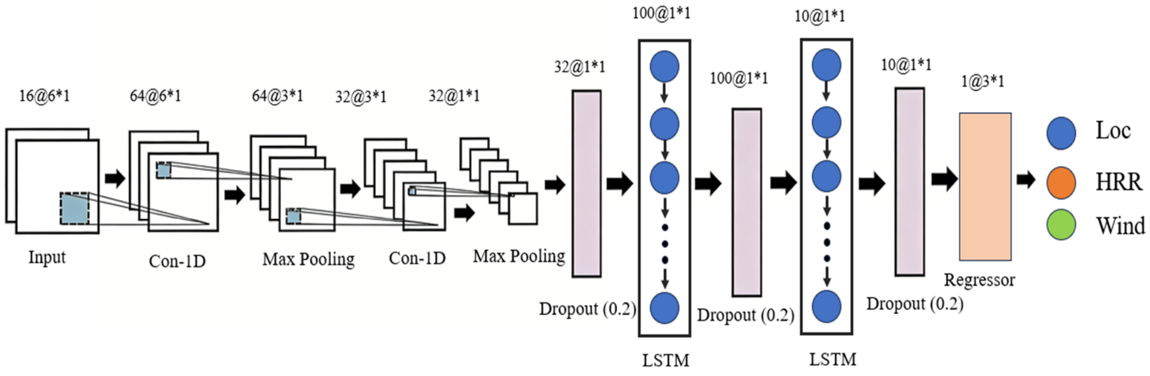
CNN-BiLSTM: Replaces LSTM with BiLSTM based on CNN-LSTM, simultaneously strengthening spatial feature extraction and bidirectional temporal modeling. Figure 5 illustrates shows the specific network structures of five representative models under conditions of 10 m sensor spacing and a 6 s observation window.



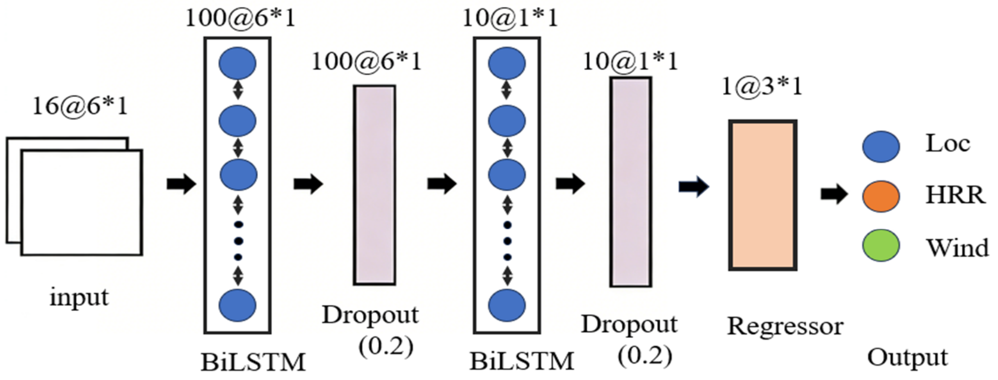
(a) Structure of BPNN Model



(b) Structure of LSTM Model

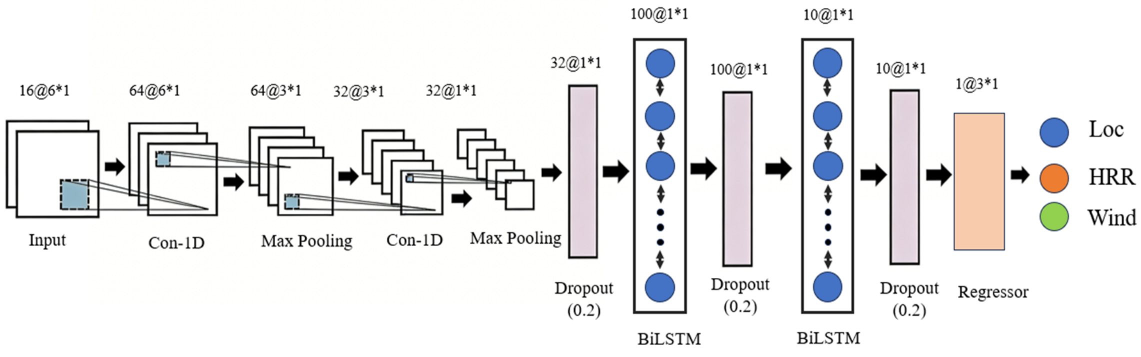


(c) Structure of CNN-LSTM Model



(d) Structure of BiLSTM Model

doi:10.20944/preprints202604.0757.v1



(e) Structure of CNN-BiLSTM

Figure 5. Schematic Diagram of BPNN, LSTM, CNN-LSTM, BiILSTM, CNN-BiLSTM Structures.

2.3.2. Hyperparameter Settings

To ensure experimental fairness and reproducibility, all models employed unified training hyperparameter configurations. The optimizer was set as RMSprop [25], with an initial learning rate of 1×10^{-3}, a batch size of 800, and the number of training epochs set to 500. The loss function was Mean Squared Error (MSE). Input data were uniformly normalized (zero mean, unit variance), with normalization parameters computed solely from the training set and applied to the validation and test sets.

2.3.3. Experimental Design

To systematically evaluate this study adopts a full factorial experimental design. It encompasses three sensor spacings (10 m, 20 m, 30 m), three time window lengths (6 s, 8 s, 12 s), and five deep learning models (BPNN, LSTM, CNN-LSTM, BiLSTM, and CNN-BiLSTM), resulting in a total of 45 independent experiments, as shown in Table 2.

Table 2. Deep Learning Model Experimental Design.

Sensor Spacing (m)	Time Window (s)	Model Types	Number of Experimental Groups	Purpose
10/20/30	6/8/12	BPNN/LSTM/BiLSTM/ CNN-LSTM/ CNN-BiLSTM	45	Investigate inversion performance of each model under different spacings and windows; study performance improvement with extended window length.

Changes in sensor spacing evaluate the impact of spatial sampling density on prediction accuracy: smaller spacing provides richer information but incurs higher hardware costs. Changes in time window simulate different observation durations in the early stages of a fire, examining the effect of data volume on model performance: shorter windows allow for shorter warning lead times, which is more advantageous for fire rescue. Comparative experiments across five model types aim to reveal the strengths and weaknesses of different deep learning architectures in spatiotemporal feature extraction and temporal modeling for tunnel fires.

This experiment was conducted on a computing platform featuring an Intel(R) Xeon(R) Gold 6334 CPU @ 3.60 GHz and 48 GB RAM. The study utilized Claude-3.5 [26] as the foundational platform to enable efficient intelligent collaboration across key stages of the tunnel fire intelligent detection optimization workflow.

3. Comprehensive Evaluation Agent

3.1. Evaluation Metrics

The Comprehensive Evaluation Analysis Agent is responsible for automatically aggregating model performance metrics from all experimental configurations and presenting the analysis results through multi-dimensional visualizations [27]. This study employs the Coefficient of Determination (R^2) and Root Mean Squared Error (RMSE) metrics to evaluate model performance in multi-task continuous regression [28].

The Coefficient of Determination (R^2) measures the degree to which the model explains the target variable. It typically ranges from 0 to 1, with values closer to 1 indicating better model fit [29,30].The calculation formula (1) is:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

The Root Mean Squared Error (RMSE) reflects the absolute deviation between predicted and true values, with units consistent with the target variable. Smaller values indicate higher prediction accuracy ADDIN. The calculation formula(2) is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

where y_i is the true value for the i-th sample, $\hat{y}_i$ is the predicted value for the i-th sample, $\bar{y}$ is the mean of the true values, n is the total number of samples, SS_{res} is the residual sum of squares, and SS_{tot} is the total sum of squares. The Average R^2 is the arithmetic mean of the R^2 values for the three fire parameters (location, heat release rate, ventilation speed), used to compare the overall performance of models. The Overall RMSE is the root mean squared error calculated for all three fire parameters as a single combined entity, reflecting the model's comprehensive prediction error across the three parameters.

3.2. Comprehensive Evaluation

3.2.1. Performance of the Models at a Sensor Spacing of 10 m

Based on 45 sets of full factorial experiments, this paper systematically evaluates the performance of five deep learning models in the task of fire source characterization in tunnel fires. For example, Figure 6 presents a comparison of the average coefficient of determination (R^2) and the overall root mean square error (RMSE) of the models on the validation set under conditions with a sensor spacing of 10 m and observation windows of 6 s, 8 s, and 12 s.

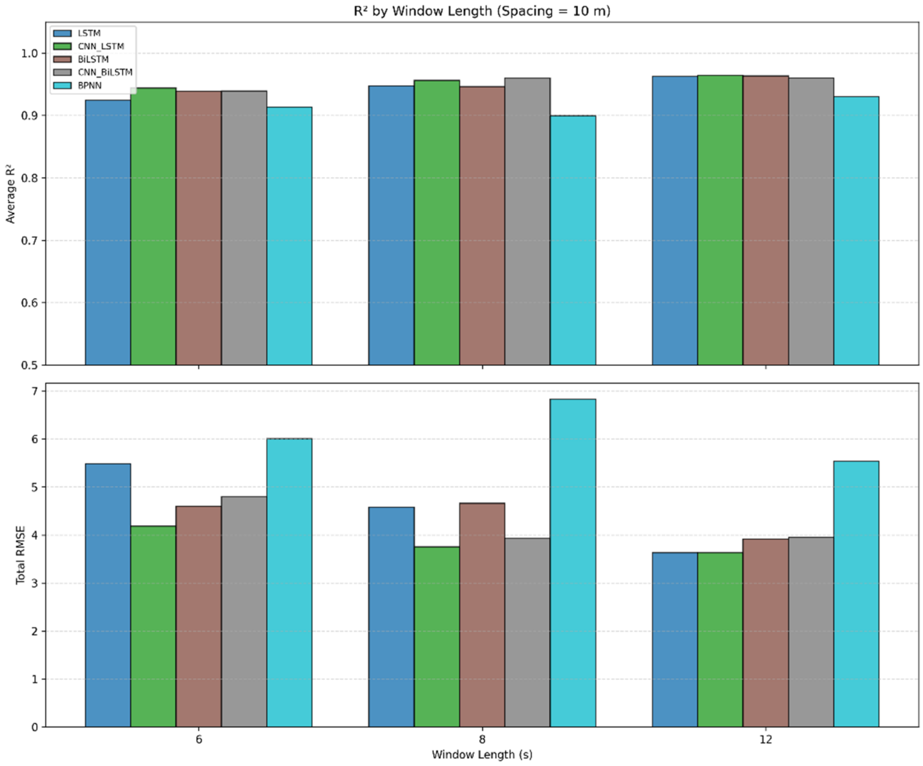


Figure 6. Key Performance Metrics of Models under 10 m Sensor Spacing and 6 s Observation Window.

From Figure 6, the following conclusions can be drawn:

(1) The performance of all temporal models significantly surpasses the BPNN model, which lacks temporal awareness. This result indicates that BPNN, by flattening the time-series data, loses critical temporal dependencies and struggles to accurately model the dynamic development of the fire.

(2) Under the shortest observation window of 6 s, the R² of CNN-LSTM (0.942) already significantly exceeds that of LSTM under the same window (0.924) and even approaches the performance of LSTM under the 8 s window (0.949). This suggests a spatial correlation exists among temperature measurement points at different positions along the tunnel ceiling. CNNs can effectively extract this spatial structural feature through 1D convolution, while LSTMs model the temporal evolution pattern. The synergy between them enables more accurate inversion of fire parameters.

(3) As the observation window extends, the R² values for all five models show an upward trend, with RMSE decreasing accordingly, indicating that more observation data help the models capture the fire development patterns more accurately.

(4) The performance of CNN-BiLSTM is slightly lower than CNN-LSTM. The introduction of the bidirectional mechanism did not yield improvement. This might be because the CNN already extracted sufficient spatial features, and the benefit of the bidirectional mechanism for short time windows was marginal, while it increased computational complexity.

3.2.2. Effect of Sensor Spacing on Model Performance

As sensor spacing increases, the R² of all models declines, while RMSE increases. This is due to the information loss resulting from decreased spatial sampling density. Under sparse deployment conditions with a spacing of 30 m and an observation window of 6 s, the validation set R² of LSTM is 0.901, which is close to the empirical threshold for engineering application (R² > 0.9 is often considered acceptable). In contrast, CNN-LSTM maintains a higher R² of 0.917, approximately 1.6% higher than LSTM. This result indicates that CNN-LSTM, through spatial feature extraction, can compensate to some extent for the information loss caused by sparse sensor deployment, potentially enabling more economical sensor placement schemes in practical engineering.

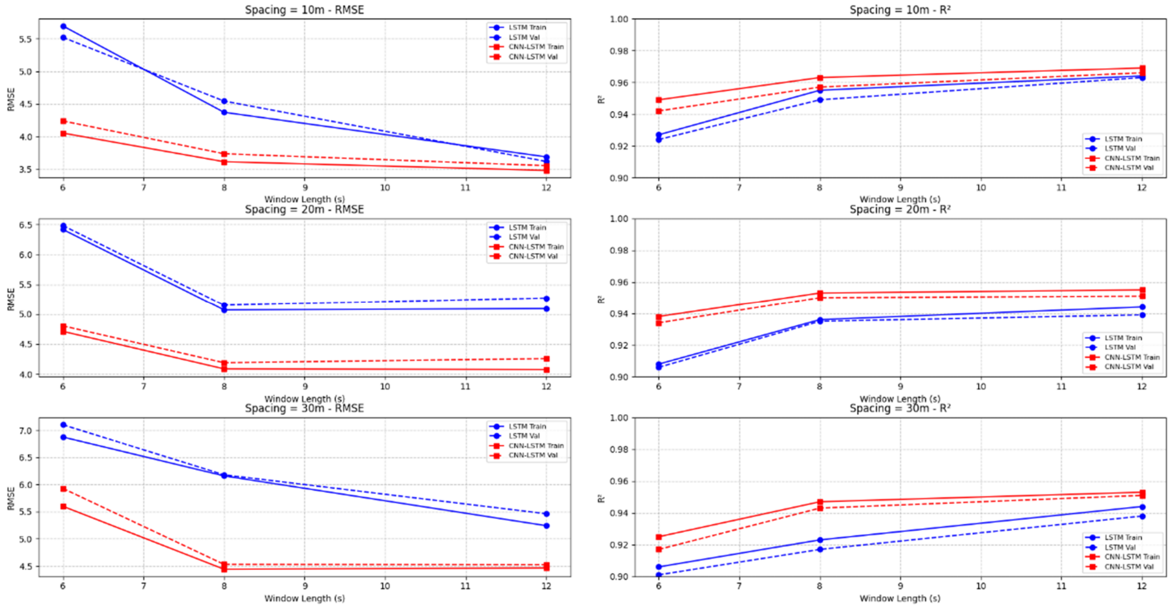


Figure 7. Key Performance of LSTM and CNN-LSTM at Sensor Spacings of 10m/20m/30m.

4. Conclusion

This paper proposed and implemented a multi-agent collaborative research framework for fire source characteristic inversion in tunnel fires. The framework decomposes the traditional tunnel fire research workflow into four core components, including physical modeling, data governance, model training, and evaluation analysis. Through a multi-agent collaboration approach based on large language models, a numerical simulation database comprising 100 tunnel fire scenarios was constructed. The performance of five deep learning models was systematically evaluated under varying sensor spacings (10 m, 20 m, 30 m) and observation windows (6 s, 8 s, 12 s), leading to the following conclusions:

(1) The multi-agent collaborative framework effectively achieves full-process automation in tunnel fire research. The four agents collaborate seamlessly, automatically completing the generation and validation of 100 FDS cases, the extraction and structuring of large scale time-series data, the training and optimization of models across 45 experimental configurations, and the multi-dimensional performance evaluation and visualization. This work compresses a traditional research workflow that would take weeks into a few hours, significantly improving research efficiency and reproducibility.

(2) The CNN-LSTM model demonstrated optimal performance in the fire source characteristic inversion task, fully validating its spatiotemporal collaborative modeling capability. Under conditions of 10 m sensor spacing and a 6 s observation window, CNN-LSTM achieved an R^2 of 0.942, which is 1.8% higher than the baseline LSTM model's R^2 of 0.924, with RMSE reduced by 28.8%, approaching the performance of the LSTM model under the 8 s window (R^2=0.949). The results indicates that the effective extraction of spatial correlations among ceiling temperature measurement points by CNN, combined with the accurate modeling of temporal evolution patterns by LSTM, achieves more precise inversion of fire parameters. Furthermore, the shorter observation window provides a valuable early warning lead time for fire rescue operations.

(3) CNN-LSTM exhibits excellent robustness under sparse sensor deployment. When sensor spacing increased to 30 m, the R^2 of LSTM dropped to a critical value of 0.901, whereas CNN-LSTM maintained a higher value of 0.917, outperforming LSTM by 1.6%. This suggests that CNN-LSTM can compensate for information loss to some extent. In practical engineering, this could enable a more economical sensor deployment scheme with 30 m spacing, reducing the number of sensors by one-third compared to the conventional 20 m spacing, thereby significantly lowering system construction costs.

This study has several limitations. Firstly, FDS simulations have inherent deviations from real fire scenarios. Secondly, the current framework focuses on simple scenarios with a single fire source and longitudinal ventilation. Thirdly, the agents lack autonomous decision-making capabilities. Future work will extend to complex scenarios such as multi-fire source coupling and incorporate advanced model architectures like attention mechanisms. In addition, validation using full-scale experimental data will be conducted, alongside the exploration of agent systems with autonomous learning capabilities, aiming to achieve a higher level of closed-loop operation toward an “AI Scientist” paradigm.

Author Contributions: Methodology, Zeng Ding, Gao Ao; software, Zeng Ding; validation, Zeng Ding and Gao Ao; formal analysis, Gao Ao and Xu Zhisheng; resources, Gao Ao; data curation, Gao Ao; writing—original draft, Zeng Ding; writing—review and editing, Zeng Ding, Gao Ao; supervision, Xu Zhisheng; project administration, Gao Ao and Xu Zhisheng. All authors have read and agreed to the published version of the manuscript.

Funding: National Science and Technology Major Project: Fire Experiment and Traceability Research Based on Full-Scale Experimental Platform，2024ZD0802105-2.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to restrictions.

Acknowledgments: During the preparation of this manuscript, we employed the Claude 3.5 multi-agent model configured with the DeepSeek-V3.2 large language model to generate FDS files, compile deep learning code, and plot analytical charts. The authors have fully reviewed, revised, and edited all AI-generated outputs, and take sole and full responsibility for the content, accuracy, and integrity of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
LLMs	Large Language Models
CFD	Computational Fluid Dynamics
FDS	Fire Dynamics Simulator
HRR	Heat Release Rate (MW)
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
CNN-LSTM	Convolutional Long Short-Term Memory
BPNN	Back Propagation Neural Network
BiLSTM	Bidirectional Long Short-Term Memory
CNN-BiLSTM	Convolutional Bidirectional Long Short-Term Memory
RMSE	Root Mean Square Error
R^2	Coefficient of Determination
RAG	Retrieval-Augmented Generation
MCP	Model Context Protocol

References

1. Zhang, X., et al., Perspectives of big experimental database and artificial intelligence in tunnel fire research. Tunnelling and Underground Space Technology, 2021. 108: p. 103691.
2. Liu, S.L., et al., Optimization of smoke exhaust efficiency under a lateral central exhaust ventilation mode in an extra-wide immersed tunnel. Journal of Zhejiang University. A. Science, 2021. 22(5): p. 396-406.

3. Zhang, T., et al., Building Artificial-Intelligence Digital Fire (AID-Fire) system: A real-scale demonstration. Journal of Building Engineering, 2022. 62: p. 105363.
4. Zhang, Y. and X. Huang, A Review of Tunnel Fire Evacuation Strategies and State-of-the-Art Research in China. Fire Technology, 2024. 60(2): p. 859-892.
5. Zhang, Z., et al., Intelligent fire location detection approach for extrawide immersed tunnels. Expert Systems with Applications, 2024. 239: p. 122251.
6. Liu, Z., et al., Research on lightweight tunnel cable fire recognition algorithm based on multi-scale features. Sci Rep, 2025. 15(1): p. 25841.
7. Hong, Y., et al., Rapid bidirectional prediction between physical field and key control parameters in tunnel fires. Tunnelling and Underground Space Technology, 2024. 153: p. 105982.
8. Xue, C.J. The road tunnel fire detection of multi-parameters based on BP neural network. in 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010) (Vol. 1, pp. 441–444). IEEE.
9. Sun, B., et al., An Improved Updatable Backpropagation Neural Network for Temperature Prognosis in Tunnel Fires. Journal of Performance of Constructed Facilities, 2022. 36(2): p. 04022012.
10. Barros-Daza, M.J., et al., Real Time Mine Fire Classification to Support Firefighter Decision Making. Fire Technology, 2022. 58(3): p. 1545-1578.
11. Guo, C., et al., Study on real-time heat release rate inversion for dynamic reconstruction and visualization of tunnel fire scenarios. Tunnelling and Underground Space Technology, 2022. 122: p. 104333.
12. Allaire, F., V. Mallet and J.B. Filippi, Emulation of wildland fire spread simulation using deep learning. Neural Networks, 2021. 141: p. 184-198.
13. Cheung, W.K., et al., Predicting smoke hazards and burning fuel via smart building fire sensor network and dual-agent deep learning. Indoor and Built Environment, 2025. 34(6): p. 1107-1125.
14. Wu, X., Y. Park and X. Huang, Smart Detection of Fire Source in Tunnel Based on the Numerical Database and Artificial Intelligence. Fire Technol, 2021. 57: p. 657-682.
15. Xie, Z.C., et al., A deep learning–based surrogate model for spatial-temporal temperature field prediction in subway tunnel fires via CFD simulation. Journal of Dynamic Disasters, 2025. 1(1): p. 100002.
16. Migoya, E., et al., A simplified model of fires in road tunnels. Comparison with three-dimensional models and full-scale measurements. Tunnelling and Underground Space Technology, 2009. 24(1): p. 37-52.
17. Zhu, Y., et al., Know Agent: Knowledge-Augmented Planning for LLM-Based Agents. 2025, Cornell University Library, arXiv.org: Ithaca. p. 3709-3732.
18. Samuel Schmidgall, Y.S.Z.W., Agent Laboratory: Using LLM Agents as Research Assistants. Empirical Methods in Natural Language Processing, 2025: p. 5977–6043.
19. Iapascurta, V., et al., Multi-Agent Approach for Sepsis Management. Healthcare informatics research, 2025. 31(2): p. 209-214.
20. Renze, M. and E. Guven, Self-reflection in llm agents: Effects on problem-solving performance. arXiv preprint arXiv:2405.06682, 2024.
21. Vermesi, I., et al., Reducing the computational requirements for simulating tunnel fires by combining multiscale modelling and multiple processor calculation. Tunnelling and Underground Space Technology, 2017. 64: p. 146-153.
22. Su, L.C., et al., Smart performance-based design for building fire safety: Prediction of smoke motion via AI. Journal of Building Engineering, 2021. 43: p. 102529.
23. Ma, Q., et al., Assist in real-time risk evaluation induced by electrical cabinet fires in nuclear power plants: A dual AI framework employing BiTCN and TCNN. Reliability Engineering & System Safety, 2025. 260: p. 111037.
24. Muksimova, S., et al., Lightweight Fire Detection in Tunnel Environments. Fire, 2025. 8(4): p. 134.
25. Qian, F., et al., Direct Prediction of the Toxic Gas Diffusion Rule in a Real Environment Based on LSTM, in International Journal of Environmental Research and Public Health. 2019. p. 2133.
26. Anthropic, S., Model card addendum: Claude 3.5 haiku and upgraded claude 3.5 sonnet. URL https://api. semanticscholar. org/CorpusID, 2024. 273639283: p. 24.

27. Wu, X., et al., An intelligent tunnel firefighting system and small-scale demonstration. Tunnelling and Underground Space Technology, 2022. 120: p. 104301.
28. Koekkoek, E.J.W. and H. Booltink, Neural network models to predict soil water retention. European Journal of Soil Science, 1999. 50(3): p. 489-495.
29. Adamowski, J. and C. Karapataki, Comparison of Multivariate Regression and Artificial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms. Journal of Hydrologic Engineering, 2010. 15(10): p. 729-743.
30. Ding, Z.; Zhang, X.; He, S.; Qi, Y.; Lin, C. Experimental and theoretical study on longitudinal deformation and internal force of shield tunnel under surcharge. Tunn. Undergr. Space Technol. 2024, 144, 105506.