Journal of Engineering, Project, and Production Management
2026, 16(2), 025-180

An Archive Resource Management System Based on Multi-Head Self-Attention Mechanism and LSTM

Hua Cui
Lecturer, School of Literature and Journalism & Communication, Yan'an University; Yan'an, 716000, China, E-mail: cuihuatougao@163.com

Project Management
Received August 24, 2025; revised October 11, 2025; accepted October 14, 2025
Available online March 7, 2026

Abstract: Archival resource management, as a critical foundation of the information society, faces challenges in efficient storage, accurate retrieval, and intelligent analysis of multimodal data. Traditional approaches relying on manual classification and keyword-based retrieval struggle with semantic understanding, predictive modeling, and real-time anomaly detection. To address these issues, this study proposes a novel archival resource management system that integrates Multi-Head Self-Attention (MHSA), Temporal Convolutional Network (TCN), and Long Short-Term Memory (LSTM). The novelty of this system lies in the combination of a global semantic representation of MHSA, multi-scale time-series modeling with TCN, and long-term dependency capture with LSTM, further enhanced by a phase mechanism and Gated Linear Unit (GLU) to optimize feature selection. Experimental evaluations on CMU and MIMIC-III datasets demonstrate that the system achieved superior performance in multiple tasks: feature similarity of 0.97, clustering accuracy of 98.25%, classification accuracy of 96.42%, F1-score of 95.28%, time-series forecasting RMSE of 0.31, MAE of 0.23, anomaly detection accuracy of 94.37%, and a low false alarm rate of 2.81%. Moreover, the system maintains robustness under noisy conditions and generalizes well across different archival datasets. These results highlight the effectiveness and originality of the proposed framework, providing a feasible solution for building intelligent archival resource management systems with high precision, predictive capability, and anomaly monitoring.

Keywords: Archival data, resource management, attention mechanism, LSTM, multi-modal data.

DOI 10.32738/JEPPM-2024-180

1. Introduction

Information technology and digital tools are rapidly developing. As a key foundation of the information society, archive resource management has become essential for organizations to accumulate knowledge, preserve history, and operate efficiently (Barry et al., 2022). Archives are not only carriers of historical information, but also important resources that support scientific decision-making, promote cultural protection, and technological innovation. In fields such as government, research institutions, and enterprises, the type, scale, and complexity of archival data are growing exponentially. Data forms are gradually expanding from traditional paper archives to multi-modal data such as electronic archives, audio and video materials, and image documents (Bozionelos et al., 2022). In this context, archive resource management systems face higher requirements, which require efficient storage and accurate retrieval, as well as intelligent archive classification, status prediction, and abnormal function monitoring (Zvarikova et al., 2022). However, the current archive resource management system exhibits many limitations when facing complex and diverse data. The traditional archive management methods mainly rely on manual classification, rule matching, and keyword-based retrieval methods. Although this type of method has a simple structure and is easy to implement, it has obvious shortcomings. Firstly, the ability to understand information is limited, making it difficult to deeply explore the semantic associations of archival data. This defect is particularly prominent in unstructured data such as text, audio, and video (Lu et al., 2023). Secondly, there is a lack of predictive ability, which makes it impossible to effectively model and predict the lifecycle and circulation trends of archives (Agarwal et al., 2022). Thirdly, insufficient real-time and intelligent capabilities result in low management efficiency when dealing with massive archive resources, and the inability to promptly detect and handle abnormal access or illegal operations (Hasanvand et al., 2023). Therefore, the research proposes an archive resource management system based on the Multi-Head Self-Attention (MHSA) mechanism and Long Short-Term Memory (LSTM), fully integrating MHSA's global understanding ability of data content with LSTM's advantages in time-series modeling. The research aims to provide a reasonable solution for building an intelligent archive resource management system with high-precision content

understanding, dynamic flow prediction, and abnormal behavior monitoring functions, and to address the shortcomings of the current system on efficiency, intelligence, and real-time response.

2. Related Works

In recent years, with the widespread application of deep learning in natural language processing and computer vision, MHSA has received widespread attention as a key technology that can capture global dependencies between data through parallelized multi-head attention computation. Researchers have conducted in-depth exploration on the application of MHSA in information understanding, classification, and sequence modeling. Vasanthi et al. (2024) addressed the visual challenges in small-scale object detection through a MHSA module to extract deep information from feature maps. The spatial pyramid pooling, a faster pooling technique, was used to aggregate feature maps at five different scales to improve the quality of feature maps, thereby enhancing the performance of small-scale object detection. Deng et al. (2022) built a novel deep learning method for predicting drug target interactions, which utilized the MHSA mechanism in deep residual networks to predict the binding affinity. This method achieved performance improvements on two benchmark datasets, surpassing multiple state-of-the-art methods currently available. Mesai Belgacem et al. (2024) proposed an innovative intelligent model that integrated gate controlled residual neural network and MHSA mechanism for effective fault recognition and diagnosis in photovoltaic arrays. By conducting fault simulation experiments on small-scale photovoltaic power grid systems, performance optimization in photovoltaic array fault diagnosis was achieved.

LSTM, as a classic variant of recurrent neural networks, is designed specifically to solve long sequence dependency problems. It has been extensively applied in time-series prediction and natural language processing. Bandani et al. (2023) built a resource estimation strategy based on a multiplication LSTM to address the frequent switching and inconsistent interfaces caused by dense deployment in 5G. By analyzing the observed delay that depends on the number of users, the performance optimization of small base station transmission was achieved. Hunt et al. (2022) proposed a hybrid prediction system that combined traditional physics methods and statistical prediction techniques to address the limitations of empirical formulas and data sparsity in traditional physics models for river flow prediction. LSTM was used to predict the flow rates of 10 river stations in different climate regions of the western United States, thereby improving prediction performance. Alvi et al. (2022) built a framework based on an LSTM model for early recognition of mild cognitive impairment. The framework used denoising, segmentation, down sampling, and an LSTM model to reveal deep hidden features, and a combined Sigmoid classifier to identify patients with mild cognitive impairment, thereby improving accuracy, sensitivity, and specificity.

The archive resource management system is an important tool for supporting information storage, retrieval, circulation, and security management. Its development has undergone a transformation from traditional manual management to digital and intelligent management. Bayu et al. (2022) combined big data technology with N-Gram, naive Bayes, and Laplacian smoothing technology to classify and analyze the emotions of Indonesian people with COVID-19 Disease for the emotional archive management. Thus, the evaluation and optimization of the population's happiness level were achieved, providing new ideas for resource archive management. Andronie et al. (2023) aimed to improve efficiency and security in financial technology archive resource management. Financial technology based on artificial intelligence and the Internet of Things was utilized to optimize archive storage, retrieval, and circulation processes through digital transformation, thereby enhancing the operational efficiency and information security of archive management. Gadde (2024) proposed a method for integrating artificial intelligence-driven algorithms into modern database management systems to improve system efficiency, query optimization, and decision-making in real-time data analysis. This method enhanced system adaptability and performance by automating routine tasks such as indexing and query execution planning.

In summary, MHSA performs well in feature extraction and global modeling, LSTM excels in time-series prediction, and archive resource management is transitioning towards intelligence. However, existing research has mostly focused on the individual applications of MHSA or LSTM, lacking collaborative exploration between the two in archive management, and there are still shortcomings in terms of performance and adaptability in large-scale multi-modal data processing. Therefore, this study proposes an archive resource management system. The innovation lies in further introducing the Temporal Convolutional Network (TCN) based on MHSA and LSTM, which combines the advantages of these three to organically combine archive resource content understanding and time-series prediction. In addition, residual connections are introduced to enhance the system’s adaptability to large-scale archival data.

3. Methods and Materials

This section provides a detailed introduction to constructing the archive resource management system. Firstly, the overall network structure of the system is described, which consists of three parts: MHSA, TCN, and LSTM. Subsequently, a phase mechanism and a Gated Linear Unit (GLU) are introduced on this basis to further optimize.

3.1. Archive Resource Management System based on MHSA and LSTM

The archive resource management system requires efficient storage, classification, retrieval, and circulation management of complex and diverse data. Traditional keyword matching and rule-based methods are no longer sufficient to meet the demand for intelligence. Especially when dealing with multi-modal data, including text, images, timestamps, etc., there are problems with insufficient feature extraction capabilities and weak modeling capabilities for time-series (Toharudin et al., 2023). Therefore, an archive resource management system based on MHSA and LSTM is proposed. To further improve the efficiency and stability of time-series modeling, TCN is introduced to construct a complete system structure, enhancing the system’s ability to process large-scale data and parallel computing performance. The proposed system structure is shown in Fig. 1.

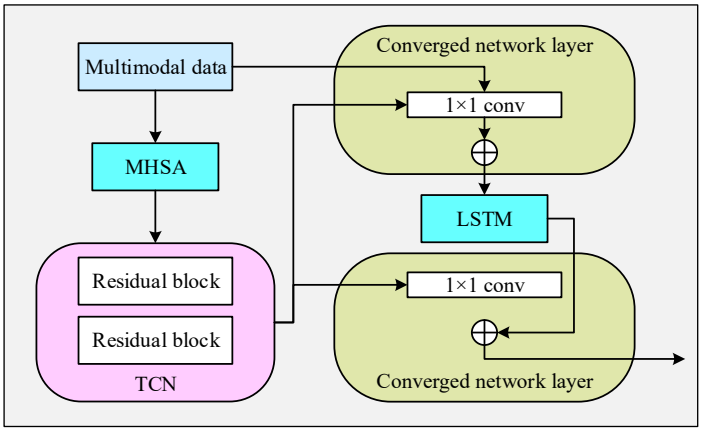


Fig. 1. Schematic diagram of network structure of archives resource management system

In Fig. 1, in addition to the three parts of MHSA, TCN, and LSTM, the system also integrates and optimizes features through two fusion network layers (1×1 convolution) and residual connections. Firstly, the input data is used to extract global dependencies through MHSA, capture important features between data, and pass the results to the TCN module for time-series modeling. TCN effectively models short-term and long-term dependencies through causal convolution and dilated convolution, and outputs efficient time-series features. Meanwhile, the results of MHSA are directly fused and dimensionality reduced with the TCN output through residual connections in the fusion network layer. Subsequently, the fused features are input into the LSTM to further capture long-term temporal dependencies in the data and optimize the modeling effect of archive circulation patterns. The output of LSTM undergoes final feature integration through the second fusion network layer, combined with residual connections to ensure the integrity of information transmission and stability. Ultimately, the system outputs data for classification, circulation prediction, and anomaly detection of archival resources, achieving efficient collaboration between global feature extraction and time-series modeling. In the system, the structure of MHSA is shown in Fig. 2.

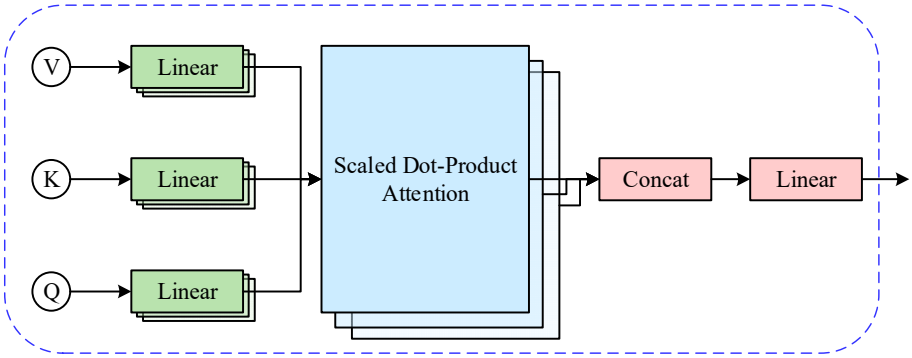


Fig. 2. Schematic diagram of MHSA structure

As shown in Fig. 2, the core of MHSA includes Query (Q), Key (K), and Value (V), where Q is dot product with K to generate attention weights, and K is used to match Q to determine the importance of each input. V is used to output the final feature representation after attention weighting. The linear layer in MHSA projects the input vector to generate Q, K, and V of multiple heads. Scaling dot product attention combines V with weighted weights by calculating attention weights to generate attention output. In the archive resource management system, input vectors can be text content, timestamp vectors, or feature representations of images. Q, K, and V are all obtained by linearly transforming the input vectors, as shown in Eq. (1).

$$Q = XW_Q, K = XW_K, V = XW_V \tag{1}$$

In Eq. (1), X represents the input vector, and W_Q, W_K and W_V signify the weight matrices of Q, K, and V vectors. In MHSA, attention weights are calculated by scaling dot product attention, as shown in Eq. (2).

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{2}$$

In Eq. (2), QK^T represents the dot product of the query vector and the key vector, resulting in an attention score matrix, which indicates the correlation between each archive data and other archive data. $\sqrt{d_k}$ represents the scaling factor to avoid the instability of softmax output caused by excessive QK^T-dot product results. In archive resource management, the similarity between Q and K determines which archives have strong correlation, such as the similarity of archive text content,

archive circulation patterns with close timestamps, and the degree of association between multi-modal data. Each head in MHSA independently learns information from different subspaces (Pansara, 2022). Ultimately, the results from multiple heads in MHSA will be concatenated, and a global feature representation will be generated via a linear transformation. The TCN structure in the system is shown in Fig. 3.

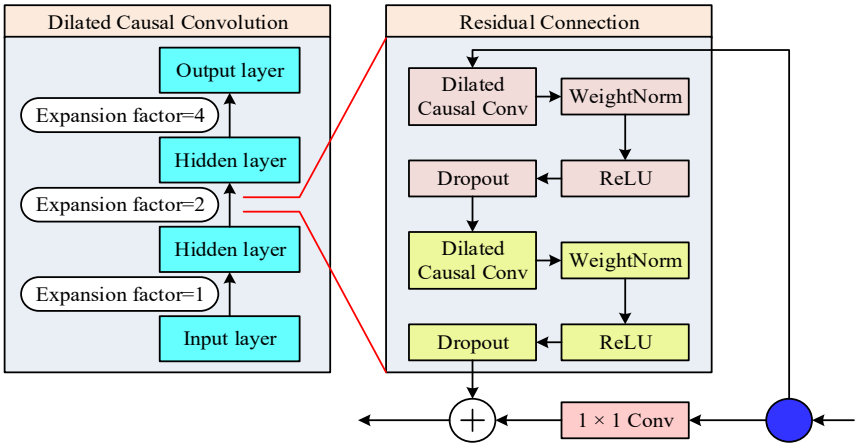


Fig. 3. Structure diagram of TCN

As shown in Fig. 3, the structure of TCN has multi-layer dilated causal convolution and residual connections. The input data is sequentially processed through multi-layer dilated convolutions. The dilation factor of each layer increases layer by layer to expand the receptive field of the convolution kernel and capture features at different time scales. Each convolutional layer contains ReLU activation function, weight normalization, and Dropout regularization to ensure the training stability and the generalization ability. In addition, residual connections directly transmit input data to the output, ensuring efficient transmission of feature information in the network. Finally, feature dimensionality reduction and fusion are achieved through 1×1 convolution.

In TCN, the expansion factor values are 1, 2, and 4, corresponding to short-term dependence, medium-term dependence, and long-term dependence. Short-term dependence can capture the patterns of recent archive borrowing and archiving, medium-term dependence can identify periodic behavior within a continuous period of time, and long-term dependence can discover archive usage patterns across months or seasons. Based on multi-layer dilated convolutions and residual block calculations, TCN generates high-dimensional feature representations of time-series that can predict borrowing or archiving trends, identify usage peaks and valleys, and capture abnormal access behavior, improving system management and security. LSTM can further model long-term temporal dependencies based on the high-dimensional features of time-series generated by TCN, dynamically adjust information flow through input gates, forget gates, and output gates, and capture complex dynamic changes and long-term trends in the circulation of archival resources. In addition, LSTM effectively integrates historical and current features, optimizes the prediction accuracy of time-series, and models archive circulation patterns.

3.2. System Optimization Design Integrating Phase Mechanism and GLU

The system proposed in the study, which combines MHSA, TCN, and LSTM, can efficiently extract global features of archival data and model short-term and long-term dependencies of time-series, providing substantial content for the classification, circulation prediction, and abnormal behavior detection of archival resources. However, when the system processes data with long time spans, there may be information flow attenuation, which affects the capture of long-term dependency relationships. Meanwhile, traditional activation functions are too simple in feature selection, failing to fully retain key information, resulting in feature redundancy and decreased computational efficiency (Logeshwaran, 2022). Therefore, further research is conducted to optimize the system by integrating phase mechanisms with temporal phase shifts and GLUs that automatically screen and preserve important information in input features through gating mechanisms. The optimized system network structure is shown in Fig. 4.

From Fig. 4, the main network structure of the optimized system has not changed, and the role of each part remains unchanged. The six key time features selected for archive access in the input data, including timestamp offset, periodic usage time feature value, time interval value of circulation events, time node offset value of status updates, time feature value of abnormal behavior detection, and multi-modal data synchronization offset, are used as the main inputs for the phase mechanism. These data represent the key information of time features in archive resource management and can effectively reflect the time dependency relationship between archive circulation and usage. In addition, GLU is added to TCN because GLU, combined with the extended causal convolution of TCN, can further optimize the efficiency and accuracy of time-series feature extraction and enhance the ability to express information at different time scales. The six main inputs of the phase mechanism are named T1-T6, and their data transmission in LSTM is shown in Fig. 5.

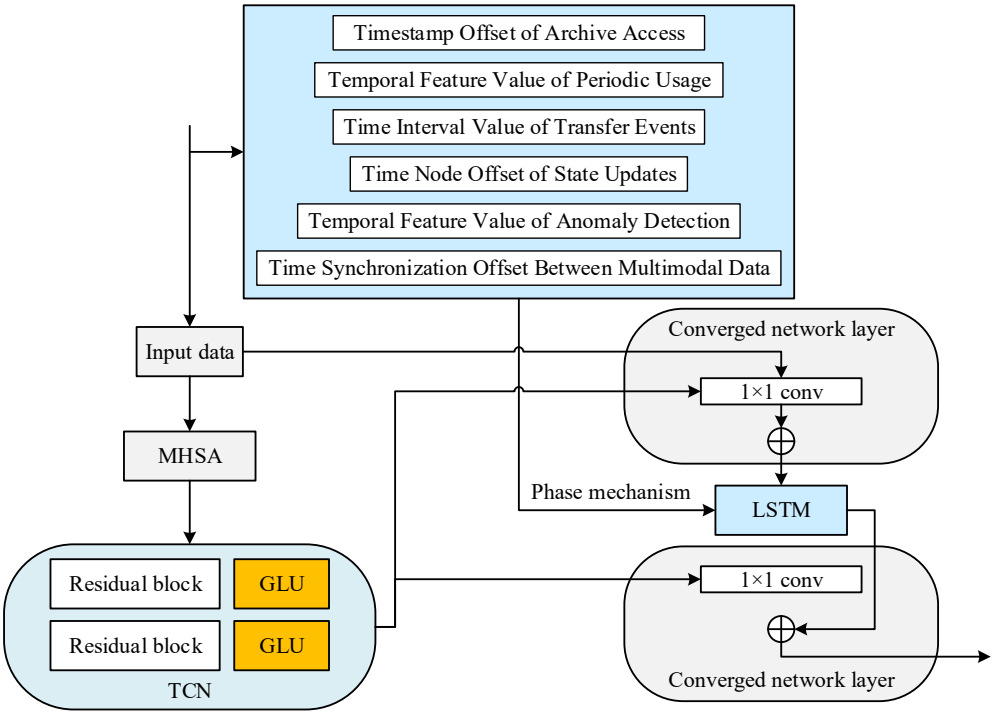


Fig. 4. Schematic diagram of the optimized system network structure

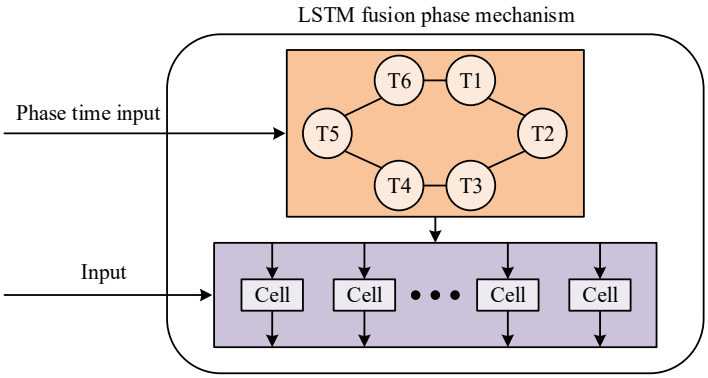


Fig. 5. Schematic diagram of LSTM fusion phase mechanism

As shown in Fig. 5, the six main inputs T1-T6 of the phase mechanism are sequentially fed into each time step of the LSTM network as key time features. In the LSTM structure, these time features are fed into the cell units of each time step in batches. Specifically, T1-T6 is gradually passed as input features to the input gate, forget gate, and output gate of LSTM, dynamically updating the hidden state and cell state. Based on the gate control mechanism, LSTM can adaptively choose to retain or discard information in time features, and pass the optimized time features back to the Cell unit of the next time step. Assuming that $\{T_1, T_2, T_3, T_4, T_5, T_6\}$ dynamically learns the offset ϕ_t through the phase mechanism and forms a new input feature $\tilde{x}_t$, the calculation is shown in Eq. (3).

$$\tilde{x}_t = \text{Concat}\left(T_1 + \phi_1, T_2 + \phi_2, T_3 + \phi_3, T_4 + \phi_4, T_5 + \phi_5, T_6 + \phi_6\right) \tag{3}$$

In the operation of LSTM, T_1 and T_3 help identify short-term offsets or flow intervals, ensuring that historical dependency information is not erroneously discarded; T_2 captures periodic time characteristics, guiding the system to identify the periodic patterns of archive circulation; T_4 ensures that the update features of state nodes are preserved in long-term dependencies; T_5 provides abnormal behavior time characteristics to help the system focus on potential abnormal time periods; and T_6 ensures synchronized data updates from different time sources by aligning the time steps in phase, reducing the impact of time misalignment on modeling. In the TCN fused with GLU, the residual module is shown in Fig. 6.

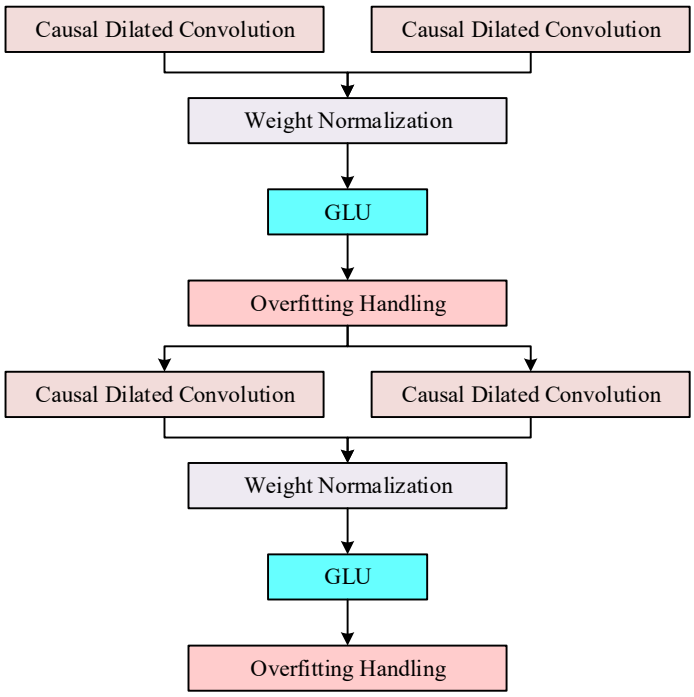


Fig. 6. Schematic diagram of residual module after GLU fusion

From Fig. 6, incorporating GLU into the one-dimensional convolution module of TCN effectively optimizes the processing of input features. Specifically, after the input data undergoes causal expansion convolution, GLU weights the features through a dynamic filtering mechanism, preserves important information, and suppresses the interference of irrelevant features, thereby achieving comprehensive utilization and focused attention on the input information. Meanwhile, the features selected by GLU are subjected to dimension matching through 1×1 convolution and fused with the input features through residual connections. The TCN that integrates GLU first performs causal dilation convolution on the input data, as shown in Eq. (4).

$$Z = W *_d X + b \tag{4}$$

In Eq. (4), W represents the convolutional kernel weight matrix; $*_d$ represents a convolution operation with an expansion factor; X represents the input feature data; b represents the convolution bias term; and Z Z represents the output feature map after the onvolution operation. GLU performs feature filtering and weighting, as shown in Eq. (5).

$$\mathrm{GLU}(Z) = \sigma\left(W_g *_d X + b_g\right) \odot \tanh\left(W_c *_d X + b_c\right) \tag{5}$$

In Eq. (5), W_g and W_c are the gating weights and convolution kernel weight matrices, respectively; b_g and b_c represent gating and convolution bias terms, respectively; $\sigma(\cdot)$ signifies the Sigmoid activation function; $tanh(\cdot)$ signifies the hyperbolic tangent function; and $\odot$ signifies element wise multiplication, which is used to weight the gating results with features. In summary, a complete archive resource management system based on MHSA and LSTM is proposed, which combines MHSA for global feature extraction and integration, TCN for modeling short-term and long-term dependencies, LSTM for in-depth modeling of dynamic temporal features, and a phase mechanism and GLU gating mechanism introduced in optimization design.

4. Results

To validate the effectiveness of the proposed archival resource management system, two representative datasets are selected: the CMU Document Understanding Dataset and the MIMIC-III clinical dataset. The former contains large-scale structured and unstructured archival resources, such as text, images, and timestamp metadata, which comprehensively simulate the complexity of multi-modal data in archival management. The latter, although originating from the medical domain, includes electronic health records of more than 40,000 patients, along with time-series event logs and multi-modal information. Its temporal dependency characteristics are highly consistent with archival circulation patterns, making it suitable as a validation platform for time-series modeling and anomaly detection. By conducting joint experiments on these two datasets, the system's performance in both multi-modal feature extraction and temporal prediction can be thoroughly evaluated, ensuring the comprehensiveness and reliability of the experimental results. To ensure reliability, the study guarantees that all experiments are conducted under the same environment and parameter settings, as displayed in Table 1.

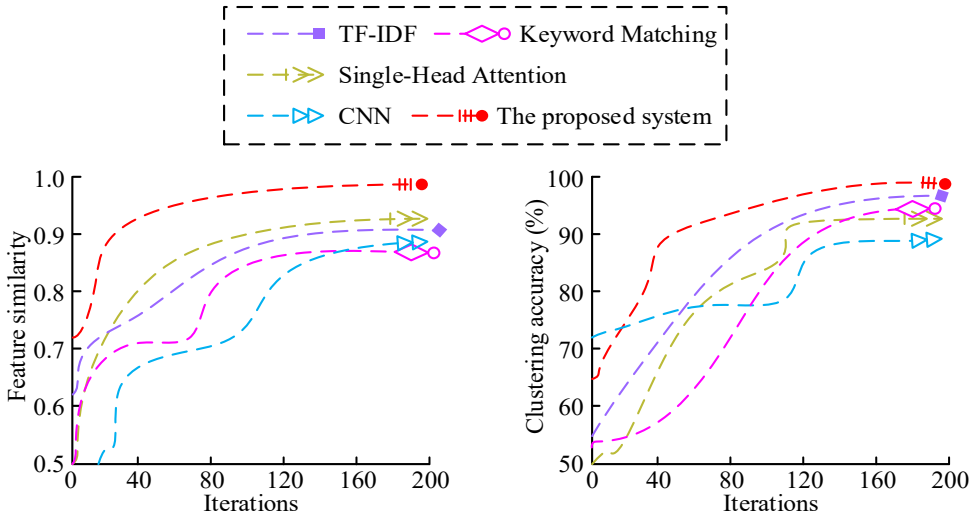
In Table 1, the parameter settings of this study were determined based on standard practices in deep learning models and further tuned through multiple preliminary experiments. Specifically, the embedding dimension of MHSA was set to 256 to balance feature representation capacity and computational efficiency. The number of attention heads was set to 8 to

achieve a trade-off between performance and resource consumption. For TCN, the kernel size was set to 3 and the model had 4 layers, capturing multi-scale temporal dependencies while controlling model depth. The LSTM was configured with 128 hidden units and 2 layers, which effectively modeled long-term dependencies while avoiding overfitting and excessive computational cost. The learning rate was set to 0.001 and the batch size to 64, both validated through experiments to accelerate convergence and maintain stability.

Based on Table 1, the study first verifies the system performance in global feature extraction. Term Frequency-Inverse Document Frequency (TF-IDF), keyword matching, single-head self-attention mechanism, and the Convolutional Neural Network (CNN) feature extraction method are selected as comparison methods. The feature similarity and clustering accuracy are taken as experimental indicators. The result is shown in Fig. 7.

Table 1. Experimental environment configuration

Environment configuration		Parameter settings	
Category	Configuration	Number of MHSA Heads	8
Processor	Intel Core i9-12900K	MHSA Embedding Dimension	256
GPU	NVIDIA RTX 3090	TCN Kernel Size	3
Memory	64GB DDR5	TCN Layers	4
Storage	2TB NVMe SSD	Dilation Factors	[1, 2, 4]
Operating System	Ubuntu 20.04 LTS	Dropout Rate	0.2
Deep Learning Framework	PyTorch 1.12 / TensorFlow 2.8	LSTM Hidden Units	128
Programming Language	Python 3.8	LSTM Layers	2
Database	MySQL 8.0 / MongoDB 5.0	Learning Rate	0.001
/	/	Batch Size	64



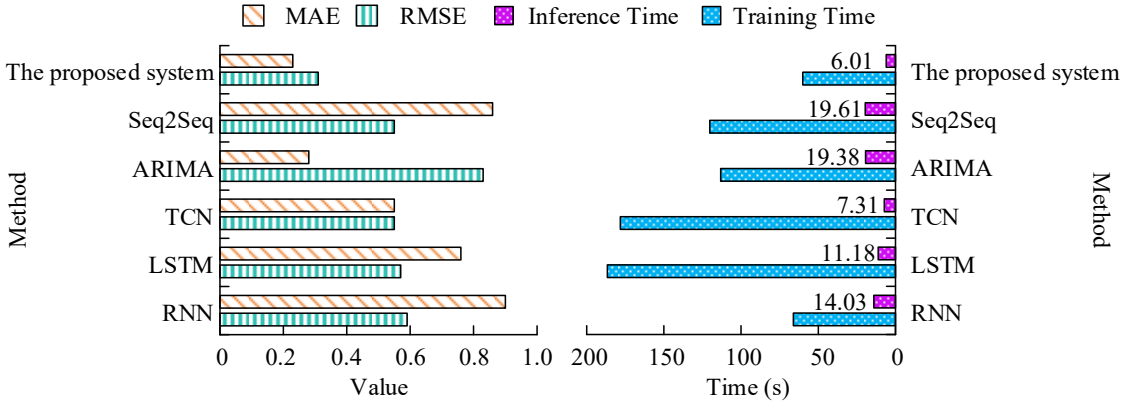
(a) Comparison of feature similarity (b) Comparison of clustering accuracy

Fig. 7. Comparison results of global feature extraction performance

From Fig. 7 (a), for feature similarity, the proposed system rapidly increased to above 0.90 in the early stages of iteration and reached 0.97 at the end of iteration. Among the four comparison methods, only the feature similarity of the single-head attention mechanism ultimately reached 0.90 or above. The final similarity of TF-IDF, keyword matching, and CNN was 0.89, 0.86, and 0.87, respectively. As shown in Fig. 7(b), for clustering accuracy, TF-IDF, keyword matching, the single-head attention mechanism, and CNN accounted for 96.17%, 94.02%, 92.31%, and 89.64%, respectively. The clustering accuracy of the proposed system reached 98.25% after the iteration was completed. This indicates that the proposed system can efficiently extract global features of archival data and demonstrate significant performance advantages in classification and clustering tasks. On this basis, a comparative experiment of time-series prediction is conducted, selecting the Auto-Regressive Integrated Moving Average (ARIMA), the Seq2Seq model, as well as separate LSTM and TCN models for comparison, as displayed in Fig. 8.

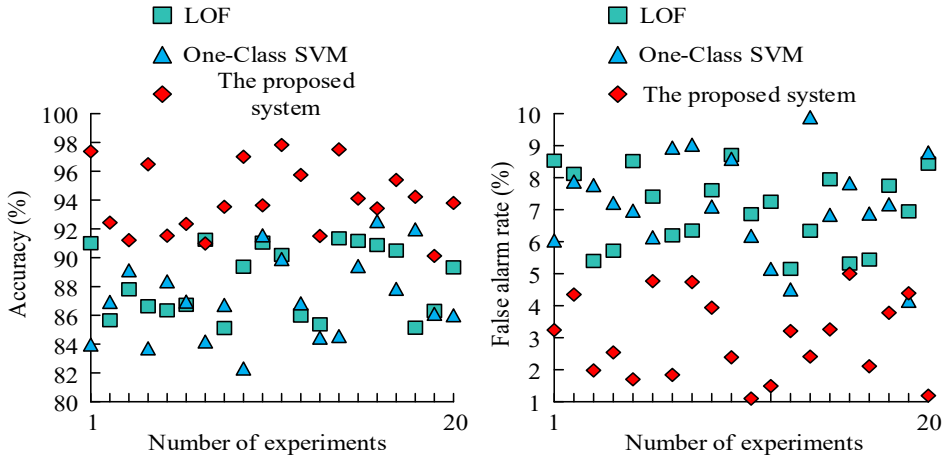
As shown in Fig. 8 (a), in terms of prediction error, the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) of the proposed system were 0.31 and 0.23, respectively, which were much lower than those of other methods. The RMSE of LSTM was 0.57, and TCN was 0.55, while the RMSE of the mixed model ARIMA and Seq2Seq were 0.83 and 0.55, respectively. As shown in Fig. 8 (b), in terms of computational efficiency, the proposed system had a training time of 62s and an inference time of 6.01s, which was superior to the LSTM with a training time of 186.62s and an inference time of 11.18s and the TCN with a training time of 178.24s and an inference time of 7.31s. Compared to ARIMA and

Seq2Seq, the proposed system significantly reduces computation time while improving accuracy. Furthermore, comparative experiments are conducted on abnormal behavior detection, selecting Local Outlier Factor (LOF) and One-Class Support Vector Machine (One-Class SVM) as comparison methods. The results are shown in Fig. 9.



(a) Comparison of prediction error (b) Comparison of computational efficiency

Fig. 8. Results of comparison experiment and time-series prediction



(a) Comparison of accuracy (b) Comparison of false alarm rate

Fig. 9. Comparison experiment of abnormal behavior detection

According to Fig. 9(a), the abnormal behavior detection accuracy for LOF fluctuated between 84% and 94%, with an average of 88.76%. The detection accuracy of One-Class SVM was not significantly different from that of LOF, averaging 87.64%. The detection accuracy of the proposed system was generally higher than that of comparison methods in each experiment, with an average of 94.37%. According to Fig. 9(b), for the false alarm rate, the LOF was 6.18%, and the One-Class SVM was 6.51%. The average false alarm rate of the proposed system was only 2.81%. The proposed system can more accurately detect abnormal behavior and effectively reduce false alarms. Furthermore, the ablation experiment is conducted in the classification and retrieval tasks to analyze the roles of various parts in the system, as displayed in Table 2.

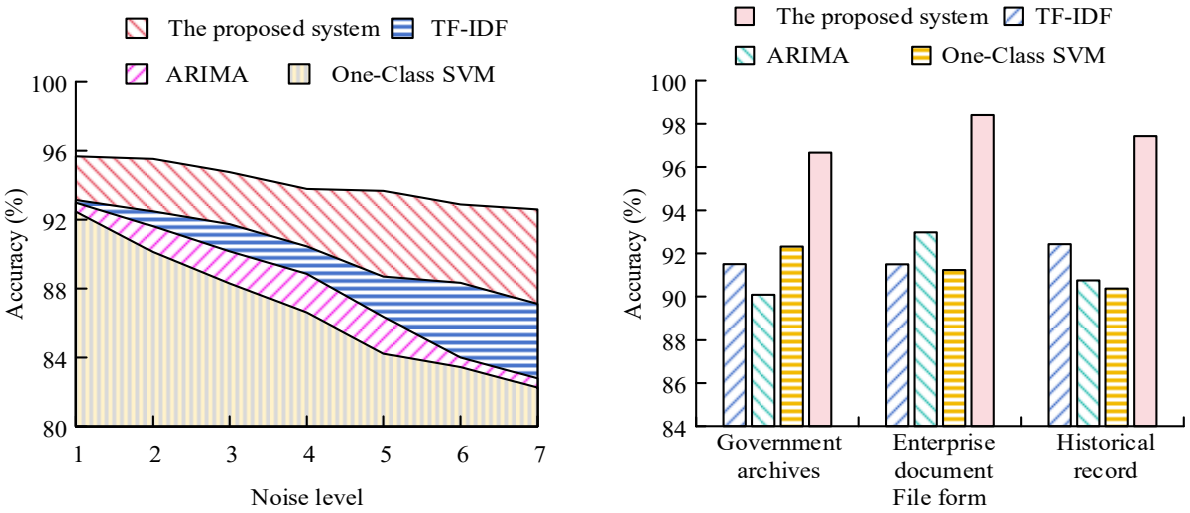
Table 2. Results of the ablation experiment

Model Configuration	Index			
	Classification Accuracy (%)	Retrieval Time (s)	Recall (%)	F1 Score (%)
The proposed system	96.42	3.12	94.15	95.28
Without GLU	92.53	4.56	90.62	91.57
Without TCN	90.47	5.31	87.53	88.97
Without MHSA	87.65	6.89	85.31	86.42
Without phase mechanism	91.12	4.98	89.23	90.15

According to Table 2, the complete system achieved 96.42% classification accuracy, 94.15% recall, 95.28% F1 score,

and 3.12 seconds of retrieval time, demonstrating the best performance. After removing GLU, the classification accuracy decreased to 92.53% and the retrieval time increased to 4.56 seconds, indicating that GLU improved the feature filtering efficiency. After removing TCN, the classification accuracy decreased to 90.47%, indicating that TCN was crucial for time-series modeling. After removing MHSA, classification accuracy decreased to 87.65% and retrieval time increased to 6.89 seconds, verifying the importance of MHSA for global feature extraction. After removing the phase mechanism, both classification accuracy and recall decreased, and the retrieval time increased to 4.98 seconds, indicating that the phase mechanism optimized time feature processing and synchronization. In summary, the collaboration among the system's modules has achieved optimal performance, verifying the rationality and effectiveness of the system. Finally, a comparative experiment is conducted on the generalization performance and robustness, and the results are shown in Fig. 10.

As shown in Fig. 10(a), with the increase of noise level, the accuracy of each method decreased. The proposed system decreased from 95.67% to 93.14%, with little change, while the TF-IDF decreased from 93.14% to 87.01%. The ARIMA decreased from 92.98% to 82.75%, and the One-Class SVM decreased from 92.46% to 82.03%. In Fig. 10(b), the accuracy of the proposed system consistently remained the highest for different types of archives, reaching 96.84% in the government archive dataset. In enterprise documents and historical records, the accuracy reached 98.36% and 97.87%, respectively. The proposed system not only has stronger generalization performance and can adapt to different types of datasets, but also exhibits good robustness in noisy environments.



(a) Results for different noise levels

(b) Results for different file forms

Fig. 10. Experimental results comparing generalization performance and robustness

5. Conclusion

An archive resource management system combining MHSA, LSTM, and TCN was proposed to address the efficiency and lack of intelligence in multi-modal data processing in archive resource management. The superior performance of the system in global feature extraction, time-series prediction, and abnormal behavior detection was verified through experiments. The experimental results showed that the proposed system exhibited significant advantages in multiple tasks. In global feature extraction, the system achieved a feature similarity of 0.97 and a clustering accuracy of 98.25%, both of which are higher than TF-IDF, keyword matching, and the single-head attention mechanism. In time-series prediction, the RMSE and MAE of the system were 0.31 and 0.23. The training time was 62 seconds, and the inference time was 6.01 seconds. The prediction accuracy and efficiency were better than LSTM, TCN, and hybrid models. In abnormal behavior detection, the system had an accuracy rate of 94.37% and a false alarm rate of only 2.81%, which was significantly better than LOF and One-Class SVM. The ablation experiment showed that the classification accuracy of the complete system was 96.42%, and the F1 Score was 95.28%. All modules played an important role in improving performance. In addition, in the generalization and robustness experiments, the accuracy of the system decreased from 95.67% to 93.14% as the noise level increased, with minimal performance fluctuations, and maintained the highest accuracy on different archive datasets, reaching up to 98.36%. In summary, the proposed system outperforms existing methods in terms of feature extraction, time modeling, and robustness.

Moreover, although the experiments are primarily conducted on public datasets, the proposed method demonstrates strong transferability. The system architecture can process multi-modal data such as text, images, and timestamps, which aligns well with the needs of archival institutions, government agencies, and enterprise document management. Robustness tests show that the accuracy remains above 93% under noisy conditions and cross-dataset evaluations, confirming its adaptability to real-world complex scenarios. In the future, the system can be applied to archival digitization, knowledge management, and multi-modal resource management, while incorporating privacy protection mechanisms to ensure compliance.

The proposed system demonstrates strong performance across multiple tasks, but certain limitations remain. Firstly, the experiments mainly rely on public datasets and do not address privacy protection issues in real-world archival management. Future work should incorporate mechanisms such as differential privacy or federated learning to ensure data security.

Secondly, the current experimental scale is relatively limited, and the system's scalability in large-scale archival repositories or distributed environments still requires further validation. Finally, practical deployment may encounter challenges such as hardware heterogeneity, complex system integration, and maintenance costs, which should be addressed in future work.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The author confirmed that no AI tools were used in the preparation of this manuscript.

References

Agarwal, B., Togou, M. A., Marco, M., and Muntean, G. M. (2022). A comprehensive survey on radio resource management in 5G HetNets: Current solutions, future trends and open issues. *IEEE Communications Surveys & Tutorials*, 24(4), 2495-2534.

Alvi, A. M., Siuly, S., and Wang, H. (2022). A long short-term memory based framework for early detection of mild cognitive impairment from EEG signals. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(2), 375-388.

Andronie, M., Iatagan, M., Uță, C., Hurloiu, I., Dijmărescu, A., and Dijmărescu, I. (2023). Big data management algorithms in artificial Internet of Things-based fintech. *Oeconomia Copernicana*, 14(3), 769-793.

Bandani, A. K., Riyazuddien, S., Bidare Divakarachari, P., Patil, S. N., and Arvind Kumar, G. (2023). Multiplicative long short-term memory-based software-defined networking for handover management in 5G network. *Signal, Image and Video Processing*, 17(6), 2933-2941.

Barry, M. and Wilkinson, A. (2022). Employee voice, psychologisation and human resource management (HRM). *Human Resource Management Journal*, 32(3), 631-646.

Bayu, I. N. K., Suarjaya, I. M. A. D., and Buana, P. W. (2022). Classification of Indonesian population's level happiness on twitter data using N-Gram, Naïve Bayes, and big data technology. *International Journal on Advanced Science, Engineering and Information Technology*, 12(5), 1944-1949.

Bozionelos, N. and Simmering, M. J. (2022). Methodological threat or myth? Evaluating the current state of evidence on common method variance in human resource management research. *Human Resource Management Journal*, 32(1), 194-215.

Deng, L., Zeng, Y., Liu, H., Liu, Z., and Liu, X. (2022). DeepMHADTA: prediction of drug-target binding affinity using multi-head self-attention and convolutional neural network. *Current Issues in Molecular Biology*, 44(5), 2287-2299.

Gadde, H. (2024). AI-augmented database management systems for real-time data analytics. *Revista de Inteligencia Artificial en Medicina*, 15(1), 616-649.

Hasanvand, M., Nooshyar, M., Moharamkhani, E., and Selyari, A. (2023, April). Machine learning methodology for identifying vehicles using image processing. In *Artificial Intelligence and Applications* (Vol. 1, No. 3, pp. 154-162).

Hunt, K. M., Matthews, G. R., Pappenberger, F., and Prudhomme, C. (2022). Using a long short-term memory (LSTM) neural network to boost river streamflow forecasts over the western United States. *Hydrology and Earth System Sciences*, 26(21), 5449-5472.

Logeshwaran, J. (2022). The control and communication management for ultra dense cloud system using fast Fourier algorithm. *ICTACT Journal on Data Science and Machine Learning*, 3(2), 281-284.

Lu, Y., Zhang, M. M., Yang, M. M., and Wang, Y. (2023). Sustainable human resource management practices, employee resilience, and employee outcomes: Toward common good values. *Human Resource Management*, 62(3), 331-353.

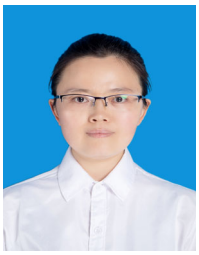
Mesai Belgacem, A., Hadef, M., and Djerdir, A. (2024). Fault diagnosis of photovoltaic array based on gated residual network with multi-head self attention mechanism. *Turkish Journal of Electrical Engineering and Computer Sciences*, 32(6), 806-828.

Pansara, R. R. (2022). Edge computing in master data management: Enhancing data processing at the source. *International Transactions in Artificial Intelligence*, 6(6), 1-11.

Toharudin, T., Pontoh, R. S., Caraka, R. E., Zahroh, S., Lee, Y., and Chen, R. C. (2023). Employing long short-term memory and Facebook prophet model in air temperature forecasting. *Communications in Statistics-Simulation and Computation*, 52(2), 279-290.

Vasanthi, P. and Mohan, L. (2024). Multi-Head-Self-Attention based YOLOv5X-transformer for multi-scale object detection. *Multimedia Tools and Applications*, 83(12), 36491-36517.

Zvarikova, K., Cug, J., and Hamilton, S. (2022). Virtual human resource management in the metaverse: Immersive work environments, data visualization tools and algorithms, and behavioral analytics. *Psychosociological Issues in Human Resource Management*, 10(1), 7-20.



Hua Cui obtained her master's degree in Management from Yan'an University in 2010. She is currently a full-time teacher at the School of Literature and Journalism Communication of Yan'an University. Her research area includes document and archival science.