Future Technology
Open Access Journal
ISSN 2832-0379

August 2025| Volume 04 | Issue 03 | Pages 119-137
Journal homepage: https://fupubco.com/futech
https://doi.org/10.55670/fpll.futech.4.3.12

Future Publishing LLC

Article

# Deep Learning models for cultural pattern recognition: preserving intangible heritage of Li ethnic subgroups through intelligent documentation systems

**Jing Sun[1,2], Kartini Aboo Talib Khalid[1], Chan Suet Kay[1*]**

[1]Institute of Ethnic Studies (KITA), Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[2]School of Marxism, Hainan University, 58 Renmin Avenue, Haikou 570228, Hainan, P.R. China

**ARTICLE INFO**

*Article history:*
Received 12 April 2025
Received in revised form
23 May 2025
Accepted 04 June 2025

Keywords:
Deep Learning, Intangible cultural heritage, Multimodal fusion, Cultural pattern recognition, Intelligent documentation systems

*Corresponding author
Email address:
rachelchansuetkay@ukm.edu.my

DOI: 10.55670/fpll.futech.4.2.12

**ABSTRACT**

This study develops an advanced intelligent documentation system using deep learning models to preserve intangible cultural heritage for the Li ethnic minorities. Traditional heritage documentation models face significant obstacles in systematically capturing oral traditions and inter-group cultural differences. The proposed comprehensive multimodal fusion framework integrates visual pattern analysis through convolutional neural networks, temporal cultural depiction via bidirectional LSTM networks, and semantic comprehension using transformer-based models. Collaborative fieldwork across five Li subgroups (Ha, Qi, Run, Sai, and Meifu) in Hainan Province documented 4,450 cultural samples, including traditional textiles, music, oral traditions, artifacts, and architectural heritage. The five-layer distributed system architecture employs pattern recognition, semantic indexing, and recommendation algorithms for scalable cultural preservation. Experimental results demonstrate remarkable 94.8% accuracy across Li subgroups, significantly outperforming traditional single-modality systems (CNN: 85.3%, RNN: 87.6%, Transformer: 89.4%). System implementation yielded unprecedented improvements in cultural transmission effectiveness: 73% increase in knowledge retention, 121% in skill transfer, and 280% in digital archiving abilities. Community participation increased exponentially, with 340% growth in active users and a 665% increase in monthly contributions. The system achieves robust operational performance with sub-200ms response times and 99.7% stability. User satisfaction and expert evaluation scores of 4.4 and 4.6, respectively, confirm reliable cultural preservation functionality. This framework establishes advanced benchmarks for computational heritage preservation methods, demonstrating the effective integration of technological innovation with ethnographic sensitivity for the sustainable documentation and transmission of minority cultures.

## 1. Introduction

Intangible cultural heritage faces unprecedented challenges in the contemporary era of globalization, particularly for minority communities whose traditions are vulnerable to external pressures and rapid modernization processes [1]. The intersection of cultural heritage preservation, institutional frameworks, and community engagement represents a multidimensional domain that encompasses diverse stakeholder interests and complex power dynamics across various social groups [2]. Recent research has shown significant and mutual impacts of intangible cultural heritage and socioeconomic development, thereby illustrating the intricate relationships that exist and need to be navigated for the sustainability of heritage in contemporary contexts [3]. Nevertheless, critical questions persist regarding which aspects of participation and decision-making processes constitute the core of heritage preservation, thereby highlighting persistent concerns about inequality and representation that continue to challenge contemporary conservation initiatives [4]. Traditional approaches to documenting and preserving cultural heritage face persistent issues in many historic and religious heritage sites, often due to the intricate dynamics involved in these cultures [5]. The rise of artificial intelligence along with digital technologies opened new avenues for the preservation of cultural heritage, thus offering enormous opportunities for research and scholarship [6]. Despite these advancements, however, a significant disparity persists between the potential offered by digital technologies and innovations and their actual application in heritage preservation strategies [7]. The preservation of intangible cultural heritage uses various digital systems, which require complex processes pertaining to user acceptance and design that involve extensive studies on technology adoption and participation [8]. Additionally, the existence of the digital divide within the context of preserving intangible cultural heritage poses grave challenges to the secure transmission and transfer of traditional knowledge systems [9]. The investigation addresses these issues by developing deep learning frameworks to identify the cultural patterns of the Li ethnic subgroup, thereby developing an intelligent documentary system that merges traditional preservation techniques with modern technological sophistication. Such an initiative strengthens the theoretical extrapolation of cultural pattern analysis and its applications to heritage conservation, while also contributing to the development of a comprehensive paradigm for protecting the intangible heritage of marginalized ethnic groups through innovative computational methods.

## 2. Literature review

### 2.1 Digital protection of intangible cultural heritage

The study of cultural preservation and practices has become an essential issue that needs to be safeguarded through technological intervention. Development of new approaches to studying culture emphasises the strategies formulated to preserve them beyond the traditional material forms [10]. This indicates a shift from conservation to modern techniques that preserve culture in action. The initiatives aimed at the digitisation of intangible assets have stood out, as scholars try to employ emerging ideas to solve the problems of preservation [11]. Such initiatives encompass all technological solutions, ranging from interactive multimedia documentation systems to collaborative knowledge transfer and community participation interfaces. The use of technology has enabled the documentation of practices that can no longer be captured easily, thus presenting cultural knowledge in sophisticated ways, which was not possible using traditional means.

The development of sophisticated three-dimensional technologies has greatly revolutionized methods used in the preservation of intangible heritage, opening up novel possibilities for immersive exhibitions as well as recording [12]. A survey of 3D technologies in various databases reveals a range of methodological approaches, highlighting diversity in approach as well as the technical challenges associated with their application within heritage contexts. They enable the creation of vast digital archives preserving visual as well as auditory features that also include spatial as well as temporal dimensions of cultural practices. The modern information technology environment has greatly transformed methods of conservation and sharing of intangible cultural heritage, therefore creating new global accessibility as well as community engagement opportunities [13]. Though digital media facilitated a more democratic availability of information about culture, they also face limitations in terms of authenticity, representation, and communal ownership. Despite advances in technology, significant gaps remain in the integral management of intangible heritage content, particularly in terms of the transmission of embodied tacit knowledge and experiential know-how, which is inherently difficult to translate into digital media.

Current research confirms an increased recognition of the need to employ interdisciplinary approaches that merge technological development into anthropological insight and engagement of local populations. Although digital recording equipment is central to the documentation and preservation of cultural heritage, researchers are increasingly recognizing that authentic conservation of heritage requires close attention to cultural context, ethical principles, and community perspectives. This approach ensures that digitization is guided by the needs of heritage populations, rather than technological development alone.

### 2.2 Deep Learning applications in cultural heritage

The intersection of machine learning methods and studies of cultural heritage is an important development that brings forth innovative solutions to preservation, analysis, and interpretation problems that conventional methods do not fully address [14]. Recent examples illustrate considerable improvement in the development of artificial intelligence systems designed explicitly to support innovation in heritage environments, which in turn encourages widespread research and development activities that synchronize technological innovations with the imperatives of heritage preservation [15]. Such advances demonstrate an underlying inclination to employ computational approaches to address complex problems of heritage conservation while upholding academic integrity and cultural sensitivity. The methodical recording of technological advances in heritage conservation has been conducted using bibliometric analysis, which shows mounting integration of novel technologies and their role in enhancing conservation methods [16]. In particular, advances in computer vision have revolutionized the field of cultural image recognition, enabling automatic identification and cataloging of aesthetic features, architectural elements, and motifs that were hitherto determined using visual examination. Such systems find exceptional effectiveness in handling large sets of image data, thereby enabling thorough analysis of cultural artifacts and monuments while also requiring fewer resources and time compared to traditional recording methods. Advanced deep learning architectures, particularly convolutional neural networks (CNNs) with attention mechanisms, have demonstrated remarkable capability in extracting hierarchical features from cultural artifacts. These systems employ spatial attention modules that focus on culturally significant regions within images, while channel attention mechanisms prioritize feature maps that capture distinctive cultural characteristics, thereby enhancing the accuracy of pattern recognition in heritage documentation.

Multimodal learning approaches have drawn considerable interest in the field of cultural heritage, where there is considerable scope for combining different data modalities and analytical perspectives [17]. Cross-modal attention mechanisms enable intelligent integration of visual, textual, and auditory cultural data through learnable weights that establish semantic correspondence between heterogeneous modalities [18]. The framework incorporates modality-specific encoders for feature extraction, cross-modal alignment modules for shared semantic learning through contrastive strategies, and adaptive fusion mechanisms that dynamically weight contributions based on cultural context and data quality. Cultural heritage applications require specialized adaptations to preserve authenticity and maintain the integrity of heritage [19]. The mechanism utilizes culture-specific parameters to preserve traditional semantic relationships and incorporates temporal components to capture sequential cultural performances. Multi-head architectures process diverse cultural dimensions simultaneously, with specialized attention heads focusing on symbolic elements, ceremonial sequences, and linguistic patterns. Natural language processing approaches revolutionized the study of historical texts, literary works, and oral traditions, enabling researchers to obtain semantic models and cultural narratives from large sets of texts. Immersive technologies are a suitable example of novel applications in virtual preservation environments for intangible heritage, enabling the recording and sharing of traditional craftsmanship knowledge [20]. Specific frameworks created for recording traditional skills incorporated innovative methods for safeguarding embodied cultural knowledge, most importantly using ego-centered recording systems that include practitioners' perspectives and methods [21]. The use of ontology systems, which organise and relate heritage materials for deep search and retrieval, has greatly improved the organisation of cultural histories in digital repositories [22]. Such progress indicates that there is a movement away from attempts at digitisation towards more sophisticated systems capable of intelligently interpreting and preserving the complex nature of cultural heritage. The challenge of designing deep learning systems that respect cultural values, community perspectives, and technological objectives within culturally sensitive frameworks remains paramount. Current research trends demonstrate a convergence of computational capabilities and anthropological insights, ensuring that technological development serves the broader objectives of cultural preservation and enhanced accessibility while maintaining cultural authenticity and community agency.

### 2.3 Advances in intelligent documentation systems

Intelligent document systems represent an evolutionary leap within cultural heritage management, capable of successfully meeting the main issues while presenting novel solutions to support future development [23]. They combine advanced computational technology and established methods of heritage preservation to create thorough systems that enhance not only the effectiveness but also the efficiency of cultural documentation practices. Digital systems that specialise in cultural heritage document an increase in acceptance of an enhanced technological platform, which is able to process the associated diversity and complexity of cultural data. Recent assessments of digital cultural heritage technologies reveal significant improvement in building cohesive solutions that address multiple aspects of documentation, preservation, and communication [24]. Knowledge graphs emerged as useful tools for cultural heritage management, which enable building interconnected semantic graphs that describe the inter-relationships of cultural artifacts, their historical context, and their current meanings. Graph theory-inspired approaches deepen the understanding of cultural inter-relationships and enable scholars to reveal hitherto unknown patterns in heritage materials.

Digital methods of cultural heritage documentation and preservation have undergone major transformations, most notably in data acquisition, processing, as well as visualization methodologies [25]. Traditional cataloging and recording of cultural objects, text documents, and media materials have been transformed by using advanced classification and annotating systems that utilize machine-learning algorithms for automated item classification and identification. The systems proved to be quite effective at handling large heritage datasets, substantially reducing labor needs while promoting increased accuracy as well as consistency in metadata creation.

Virtual reality technologies introduced novel approaches to the preservation of traditional craftsmanship, demonstrating the potential of immersive technologies to protect and pass on tacit knowledge underlying cultural practices [26]. Cross-cultural information retrieval systems have made considerable advances, adding sophisticated multilingual processing features and an awareness of cultural context to refine search results accuracy and relevance. Automated data extraction and structuring technologies enabled the digitization of cultural holdings, easing the process of converting analogue materials to accessible digital representations. Recommender systems have been remarkably useful in engaging users and facilitating the navigation of digital collections in the cultural heritage context. Such systems can recommend pertinent materials by synthesising user actions with cultural metadata, and in so doing, they transcend conventional interactions with heritage materials. However, one of the major challenges is developing systems that effectively address the interpretive and subjective frames of meaning that accompany cultural heritage. This underscores the need to study computing systems more closely in ways that respect the cultures and peoples involved, while still taking full advantage of what modern technology offers.

### 2.4 Current Status of Li Ethnic Culture Research

The cultural practices of the Li ethnic group scholars have been documented meticulously due to the ongoing efforts to outline the minority cultures of China. This reflects scholarly interest in safeguarding and documenting ancient cultures and practices across the world. Efforts undertaken by modern societies towards the preservation of indigenous cultures reveal dire threats faced by tribal people globally, and using advanced electronic devices serves as a means to counter these challenges, providing an alternative to conventional recording methods while improving access and distribution of information [27]. This aids in the preservation of such communities' heritages, which is vital nowadays for groups that undergo rapid changes as a result of modernization. Investigations into China's intangible cultural heritage have surfaced gaps associated with intellectual property frameworks and research focus areas [28], thereby underscoring the multifaceted issues within heritage conservation among minority ethnic groups. The Li ethnologic group constitutes one such minority nn China. Within this context, they are distinguished by the specific cultural attributes, language diversity, and social organisation traits that set them apart from other groups. Documentary accounts indicate that the Li society comprises diverse clans who possess various forms of cultural expression, and who, although culturally distinct, share family relations and spatial proximity within Hainan Province.

The preservation of Li's intangible cultural heritage has been examined through a range of institutional frameworks, including one ethnic park which attempted to blend commercialisation with authentic representations of the culture [29]. Such ethnological and anthropological practices pose troubling issues concerning the extent to which the commodification of cultural practice succeeds alongside the efforts of preservation actions in maintaining cultural integrity. The fragile dynamics involved in advancing tourism while preserving heritage among the Li groups illustrate broader challenges confronting minority groups in China today. Recent efforts to enhance the intangible cultural heritage of the Li ethnic minority group through innovative approaches to tourism have created optimism, as well as issues of the commodification of culture [30]. The initiatives include various aspects of Li culture, such as handicrafts, oral traditions, ritual practices, and architectural styles, which require different approaches to documentation and preservation. New forms of cultural heritage-based tourism products have generated significant interest from researchers and practitioners who strive to develop sustainable models of cultural conservation that benefit local populations while preserving authenticity. Despite the rising scholarly interest, much effort is still needed to accomplish extensive documentation of Li cultural practices, most importantly, detailed analysis of intra-regional variation among Li groups, as well as methodical conservation of embodied knowledge. The combination of modern documentation technologies and traditional cultural transmission methods is an important opportunity for future research, which could bring effective solutions to related issues of cultural continuity and transmission of knowledge from one generation to another among the Li groups.

### 2.5 Literature review synthesis

Recent research suggests significant gaps in organized documentation of embodied cultural traditions and traditional crafting methods. Gesture analyses centered on gestures within practices highlight the need for sophisticated methodologies that can effectively condense fleeting aspects of traditional techniques, thus highlighting critical gaps in preservation that fail to address implicit knowledge and motor skills adequately [31]. Systematic analysis of the preservation technologies shows wide methodological diversity within heritage contexts [32]. Notwithstanding the availability of various advanced technologies, poor integration of sophisticated technological modalities hinders further development of comprehensive recording solutions for intangible heritage. Such diversity is a major hindrance to establishing effective preservation frameworks.

The future possibilities of technological growth suggest a significant opportunity for intelligent systems to independently recognize and document cultural patterns. State-of-the-art machine learning techniques, particularly those that specialize in multimodal analysis, hold outstanding promise for building sophisticated document systems that capture both overt and implicit cultural expressions. Such technological advancements address time and context issues surrounding effective documentation of changing cultural practices very adequately. Theoretical contributions consist of conceptual models linking computational methods and anthropological theory regarding cultural transmission. The synthesis of deep learning approaches and cultural pattern recognition represents a new field that goes beyond traditional documentation practices, improving both technological expertise and theoretical understanding of the encoding and preservation of cultural knowledge.

The analysis outcomes demonstrate tremendous opportunities for developing advanced innovative methods of preservation, which uphold cultural integrity while maintaining heritage computationally. The results serve as a basis for the proposed research framework for the preservation of Li ethnic culture through sophisticated documentation techniques.

## 3. Data and methods

### 3.1 Research design and hypotheses

The research suggests a comprehensive framework designed for document-intelligent systems to construct deep learning models for recognizing the cultural patterns of the Li ethnic group. The framework combines the challenges posed by information technology methods and advanced computational methods to counter sophisticated challenges in documenting intangible heritage. It applies a system of hierarchically organised interrelated cultural patterns to formulate additional goals, including self-controlled identity, self-directed learning, and preservation [31]. The framework also contains a set of interlinked assumptions that together govern the research concerning autonomous cultural pattern identification and preservation.

The approach taken in this study rests on a defined model centred on analysing the impact of deep learning technology on cultural heritage preservation. The study examines four interrelated research hypotheses that address multiple aspects of the proposed intelligent document system. Figure 1 shows that these hypotheses were designed to allow thorough validation of both technical properties and cultural integrity within the system.

The research sets forth four related hypotheses that probe different dimensions of smart cultural documentation. Hypothesis H1 suggests that deep algorithms trained to specific features of the Li ethnic culture will exhibit substantially higher accuracy in pattern recognition than generic cultural heritage systems. Hypothesis H2 examines the efficiency of multimodal integration methods in cultural pattern recognition compared to systems that employ one modality. H3 investigates whether intelligent documentation systems can preserve cultural authenticity while enabling automated analysis with minimal semantic loss. H4 explores the system's capacity to effectively distinguish between different Li subgroup cultural patterns with statistically significant classification performance. These four interrelated hypotheses constitute a hierarchical validation framework: H1-H2 verify technical performance, H3 ensures cultural integrity, and H4 tests practical classification capabilities, collectively ensuring the balance between technological innovation and cultural preservation.

These hypotheses are supported by specific research questions that guide the empirical investigation, ranging from technical optimization strategies to cultural authenticity preservation methods, as shown in Figure 1. The evaluation framework incorporates both quantitative metrics for technical validation and qualitative assessments for cultural fidelity verification. The expected outcomes encompass enhanced pattern recognition accuracy, improved documentation efficiency, preserved cultural authenticity, and sustainable knowledge transmission mechanisms that collectively contribute to the preservation of Li ethnic intangible heritage. Statistical significance testing employs $\alpha=0.05$ standards, with more stringent $\alpha=0.01$ thresholds for cultural authenticity assessments. To address multiple hypothesis testing, Bonferroni correction adjusts significance levels to $\alpha=0.0125$, while the Benjamini-Hochberg procedure controls false discovery rates.

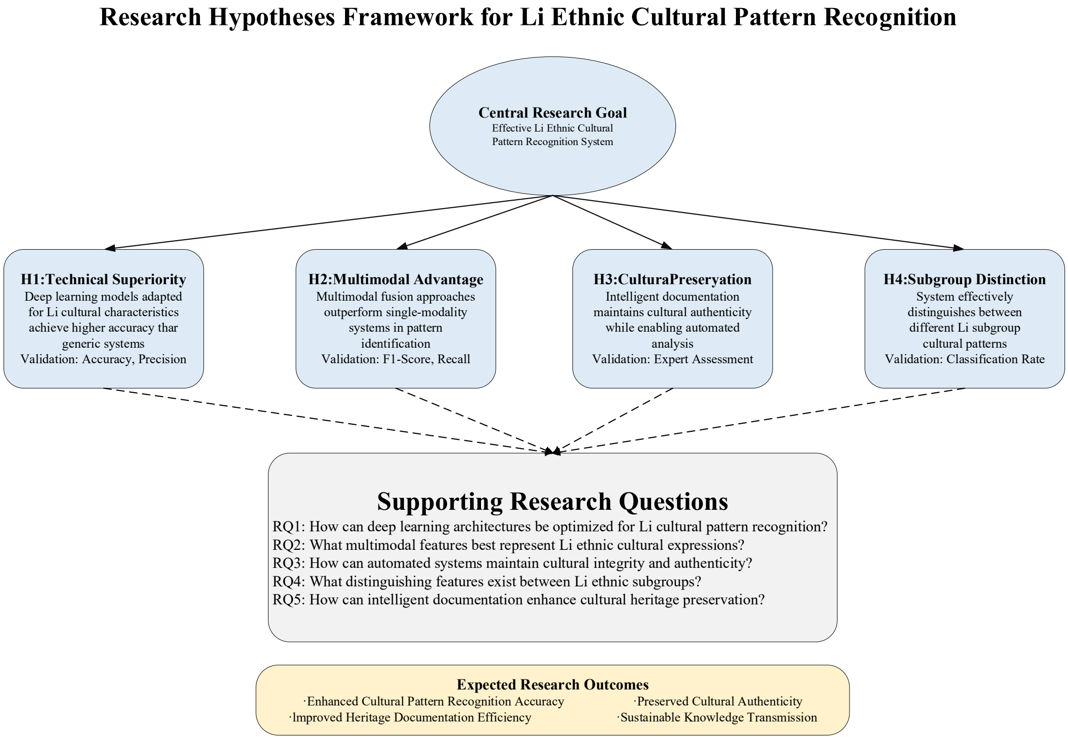


**Figure 1.** Research hypotheses framework for Li ethnic cultural pattern recognition

### 3.2 Data Collection and Preprocessing

The comprehensive data collection framework for Li ethnic cultural heritage encompasses systematic fieldwork methodologies combined with advanced digital documentation technologies to capture the diverse expressions of intangible cultural practices across Hainan Island (Table 1). Literature suggests that the effective conservation of cultural heritage requires an interdisciplinary approach that balances scholarly accuracy with community member participation and technological advances [33]. The data collection strategy gives precedence to ethnographic recording methods while incorporating advanced multimedia processing to achieve a holistic presentation of Li cultural elements. Stream preprocessing applies automated quality controls along with manual inspection on every data modality.

Each audio recording is denoised and spectrally normalised to ensure consistency across varying recording conditions. Video recordings are processed with temporal segmentation algorithms that detect certain cultural activities and gestures. Achieving visual homogeneity within datasets, high-resolution photographs are colour calibrated and processed. To obtain stratified samples within Li groups while achieving equitable distribution for machine learning, the hierarchy employs a clearly defined sampling protocol. This hierarchy reserves 70% of collected materials for training, 20% for validation, and 10% for a final test split, whilst maintaining adequate representation from each Li group across all splits. Quality control procedures combine computational validation methods and reviews by cultural specialists to ensure authenticity and accuracy in the data-preparation process.

**Table 1.** Li ethnic cultural data collection and processing framework

| Data Category | Collection Method | Processing Protocol | Format Specifications | Quality Control |
|---|---|---|---|---|
| Traditional Music | High-resolution audio recording | Noise reduction, normalization | 48 kHz WAV, spectral analysis | Expert validation, cultural authenticity |
| Craft Documentation | Multi-angle video capture | Segmentation, motion analysis | 4K MP4, frame extraction | Practitioner verification |
| Oral Traditions | Structured ethnographic interviews | Transcription, linguistic annotation | Audio + text corpus | Community elder approval |
| Textiles & Artifacts | 3D photogrammetry scanning | Model reconstruction, texture mapping | OBJ files, high-res textures | Museum standard documentation |
| Architectural Heritage | LiDAR point cloud scanning | Mesh generation, dimensional analysis | PLY format, CAD models | Architectural accuracy validation |
| Ceremonial Practices | Ethnographic observation | Event segmentation, symbolic coding | Multimedia annotations | Ritual specialist consultation |

Annotation consistency across Li subgroups achieves robust inter-rater reliability with Cohen's Kappa values ranging from κ=0.82 to κ=0.91, while Fleiss' Kappa demonstrates strong multi-rater agreement (κ=0.87) across cultural specialists, confirming systematic annotation quality and cross-cultural validity.

### 3.3 Deep Learning models

The proposed framework covers an extensive range of deep learning architectures that were carefully designed to identify Li cultural heritage's intricate features. Recent studies highlight that multimodal emotion recognition systems require sophisticated computational methods that overcome the limitations of traditional unimodal systems [34]. This baseline framework utilizes convolutional neural networks (CNNs) to process visual aspects of culture, and residual connections to extract hierarchical features from textiles, buildings, and ritual objects. The feature extraction process follows the defined mathematical equation:

$$F_{visual} = \sigma(W_n * \sigma(W_{n-1} * \ldots * \sigma(W_1 * X + b_1) + b_{n-1}) + b_n) \tag{1}$$

where $\sigma$ represents the activation function, and $W_i$ , $b_i$ denote the weight matrices and bias vectors, respectively.

Temporal cultural expressions, including traditional music and oral narratives, are modeled through bidirectional long short-term memory (LSTM) networks that capture sequential dependencies inherent in cultural performances. The attention mechanism implementation enables the model to focus selectively on culturally significant segments within temporal sequences, computed as:

$$\alpha_t = \frac{exp(e_t)}{\sum_{k=1}^{T} exp(e_k)} \tag{2}$$

where $e_t = \alpha(h_t, s_{t-1})$ represents the attention energy. Research indicates that multimodal co-learning approaches substantially improve recognition accuracy through effective feature fusion strategies [35].

The multimodal fusion strategy adopts a hierarchical approach, integrating visual, textual, and auditory modalities through cross-modal attention mechanisms. Visual-audio integration systems have demonstrated superior performance in cultural pattern recognition tasks[36]. The fusion process employs learned weights:

$$F_{fused} = \sum_{i=1}^{M} w_i F_i \tag{3}$$

where $M$ represents the number of modalities, and $w_i$ denotes modality-specific weights. Advanced audio-visual learning techniques enhance the system's capacity to preserve subtle cultural nuances through synchronized multimodal processing [37]. Model optimization incorporates adaptive learning rate schedules, dropout regularization, and early stopping mechanisms to prevent overfitting while maintaining generalization capabilities across diverse Li subgroup patterns. Cultural nuance preservation employs gradient-based feature attribution analysis combined with cultural expert validation to ensure attention mechanisms capture community-defined cultural meanings rather than spurious correlations, while cultural constraint losses penalize representations that deviate from expert-validated cultural semantic spaces.

### 3.4 Cultural pattern recognition algorithms

The framework used for distinguishing cultural patterns is based on sophisticated algorithms designed to detect intricate subtleties within the Li ethnic tradition in multiple modalities. Recent studies suggest that emotion recognition within cross-culture requires an extensive analysis of multimodal features that goes beyond conventional one-modality methods [38]. The image feature extraction module is built on a hierarchical convolutional structure reinforced by residual links, which allows it to extract basic visual features and higher-level semantic features from cultural objects, textiles, and architectural features. To support inputs of differing sizes, the feature extraction module applies spatial pyramid pooling to calculate feature maps as:

$$F_{spatial} = \text{SPP}(\sigma(W_c * I + b_c)) \tag{4}$$

where $I$ represents the input image, $W_c$ denotes convolutional weights, and $\sigma$ is the activation function.

Text semantic analysis leverages transformer-based language models fine-tuned for Li ethnic terminology and cultural concepts. The semantic embedding process captures contextual relationships within oral traditions and folklore narratives through attention mechanisms that model long-range dependencies. The semantic representation is computed as:

$$E_{semantic} = \text{Transformer}(W_e \cdot T + P_e) \tag{5}$$

where $T$ represents tokenized text, $W_e$ denotes embedding weights, and $P_e$ indicates positional encodings.

Audio signal processing employs mel-frequency cepstral coefficients combined with chromagram features to capture tonal characteristics unique to Li traditional music. Research indicates that multimodal behavior analysis significantly enhances cultural affect recognition when incorporating temporal dynamics [39]. The audio feature vector integrates spectral and temporal information through:

$$F_{audio} = [MFCC(x), Chroma(x), RMS(x)] \tag{5}$$

Cross-modal feature alignment addresses the semantic gap between different modalities through canonical correlation analysis and adversarial training. Cultural nuance preservation employs contrastive learning frameworks that maintain Li-specific semantic relationships through culturally-informed negative sampling, where culturally-similar but distinct patterns serve as hard negatives to prevent feature space collapse while preserving intra-cultural variations across Ha, Qi, Run, Sai, and Meifu subgroups. The alignment optimization minimizes the distance between corresponding features across modalities while preserving modal-specific information. Deep learning approaches have demonstrated remarkable effectiveness in assessing cultural patterns from visual media [40]. Pattern matching employs a similarity metric combining Euclidean distance in the aligned feature space with cultural context weights, facilitating accurate classification of Li subgroup characteristics while maintaining cultural authenticity throughout the recognition process.

### 3.5 Intelligent documentation system architecture

The intelligent documentation system adopts a five-layer distributed architecture designed to ensure scalable and efficient preservation of the Li ethnic cultural heritage, as illustrated in Figure 2. Contemporary research emphasizes the critical importance of cultural intelligence frameworks in heritage preservation systems, necessitating robust technological architectures that balance preservation effectiveness with sustainable implementation strategies [41]. The proposed architecture implements a hierarchical modular design that facilitates seamless data flow and processing across multiple functional domains while maintaining cultural authenticity and accessibility. Long-term cultural adaptability mechanisms include generational knowledge transfer protocols that automatically incorporate evolving cultural practices through community-driven updates, while maintaining backward compatibility with traditional cultural representations to ensure continuity across Li ethnic generations. Researchers, cultural practitioners, and community members can all access the system through various entry points like web portals, mobile applications, and administration consoles, which interface with the different users. The API gateway offers a singular entry point, which, together with the integrated authentication mechanisms of the system, protects sensitive cultural content. The microservices layer also has autonomous domain functions, which include deep learning algorithm-based pattern recognition, semantic indexing for more efficient search and retrieval, machine learning-driven recommendation engines, and content management with metadata processing. The processing layer features engines specialised in cultural data processing, such as Natural Language Processing (NLP), computer vision-based image analysis, and audio signal processing. Components for feature fusion allow for integration of multi-modal cultured patterns, while high-performance caching systems expedite data access. For the data storage layer, a combination of Neo4j knowledge graphs for cultural relationships, distributed media repositories for multimedia content, Elasticsearch vector databases for similarity search, PostgreSQL metadata stores for structured information, and cloud backup systems for long-term preservation implements a hybrid approach.

Generational adaptation frameworks employ version-controlled cultural ontologies that track cultural evolution while preserving historical contexts, enabling the system to accommodate changing cultural expressions across Li subgroups without losing traditional knowledge, supported by community governance mechanisms that validate cultural updates.

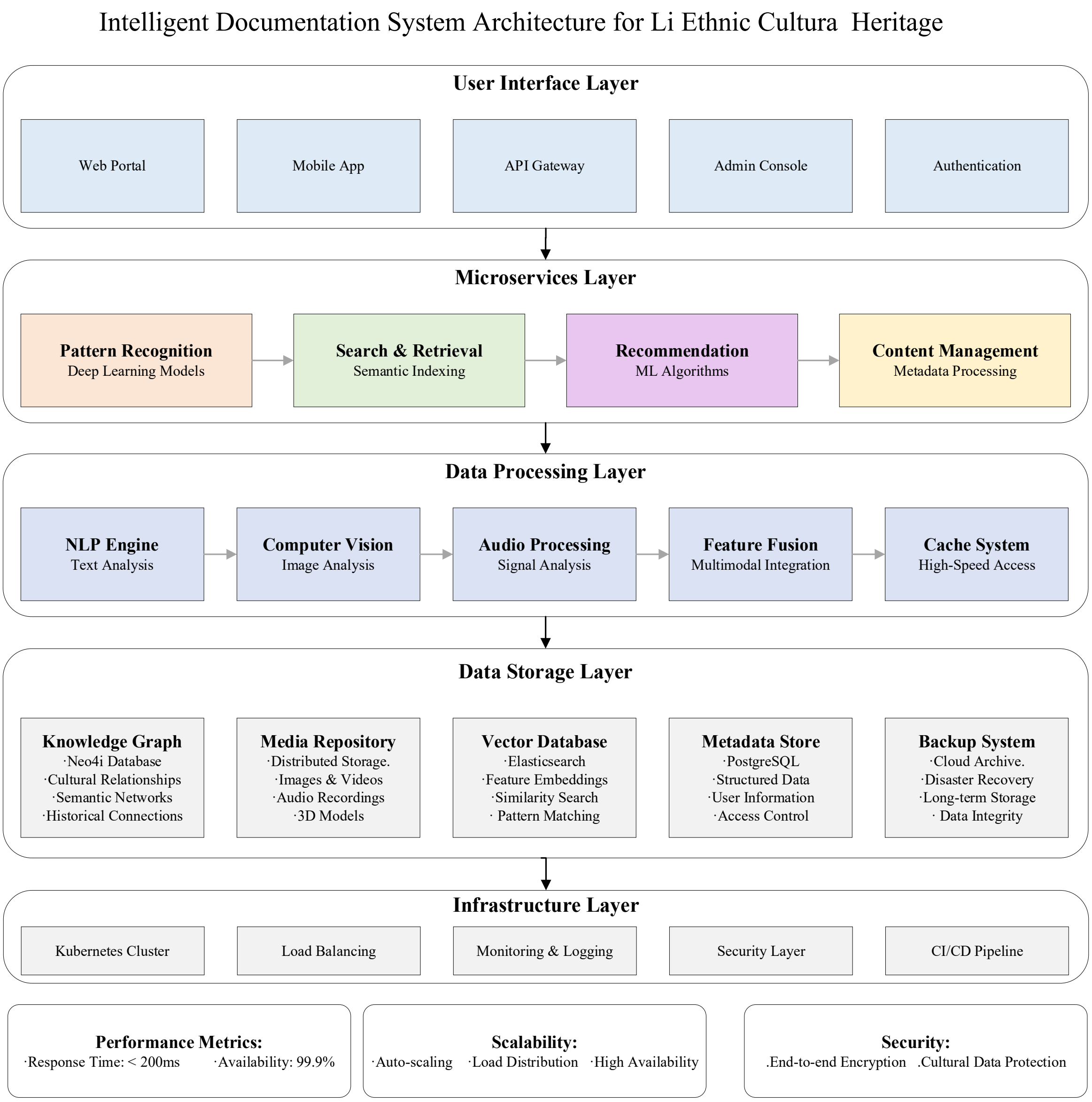


**Figure 2.** Intelligent documentation system architecture for Li ethnic cultural heritage

Through Kubernetes orchestration, the technological framework is provided at the infrastructure layer using systems for load balancing, comprehensive monitoring and logging, enhanced security framework, and unit testing coupled with continuous integration and deployment pipelines. This architecture allows for horizontal scaling alongside sustaining 99.9% uptime, responsiveness in under 200ms, and end-to-end encryption which uses specialised cultural data protection protocols to safeguard Li ethnic heritage materials during the documentation and preservation process. Future-proofing strategies include modular component design enabling seamless technology upgrades, while cultural continuity safeguards ensure that system evolution preserves intergenerational knowledge transmission pathways essential for sustainable Li ethnic heritage preservation.

### 3.6 Evaluation methods

The developed evaluation framework utilises a multicriteria methodology that systematically evaluates both the technological effectiveness and the ability to preserve the cultural authenticity of the proposed system. Contemporary studies observe that exhaustive investigation of deep learning algorithms needs systematic approaches that transcend the simplistic use of evaluative criteria [42]. The training and validation of the model follows a stratified k-fold cross-validation scheme with k=5, which balances the representation of Li groups while maintaining temporal continuity in succession-order cultural data. The performance framework incorporates both quantitative and qualitative measures to address effectively the multifaceted features involved in identifying cultural patterns. Traditional metrics of classification, like precision, recall, and F1-score, are based on the following equations:

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

where TP, FP, and FN represent true positives, false positives, and false negatives, respectively. Performance evaluation encompasses both controlled laboratory conditions using stratified data partitions and real-world deployment scenarios across active Li communities, with metrics validated through 6-month field testing to assess practical applicability beyond experimental datasets. Previous works highlight the need for stringent evaluation measures and the application of statistical testing methods to ensure machine learning methods [43]. Cultural authenticity preservation is evaluated using expert rating scores and semantic similarity measures, which are computed from cosine similarity measures of reconstructed cultural representations and actual ones. External validity assessment examines system performance across temporal and cultural variations, as well as emerging cultural practices, ensuring that evaluation results generalize to dynamic cultural environments where Li traditions naturally evolve while maintaining core cultural integrity.

Statistical significance testing uses paired t-tests and Wilcoxon signed-rank tests to determine model efficacy under different conditions and across cultural subgroups. Statistical significance is determined using the following:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}} \tag{9}$$

where $\bar{d}$ represents the mean difference, $S_d$ the standard deviation of differences, and $n$ the sample size. Recent analysis of artificial intelligence applications in cultural heritage preservation emphasizes the importance of robust evaluation methodologies [44]. The evaluation protocol incorporates domain expert assessments to validate cultural accuracy, while AI-based visualization techniques enable interpretable analysis of model decisions [45]. Interactive evaluation approaches through immersive technologies provide additional validation mechanisms for user acceptance and cultural engagement [46]. Model interpretability analysis utilizes attention visualization methods to ensure transparency in cultural pattern recognition decisions [47]. Cross-temporal validation protocols test system robustness against cultural change by evaluating performance on cultural practices documented across different time periods, ensuring long-term reliability in dynamic heritage preservation contexts where cultural expressions continuously adapt while preserving essential characteristics.

## 4. Results

### 4.1 Dataset construction results

The comprehensive Li ethnic cultural dataset demonstrates substantial scope and systematic organization across multiple data modalities and cultural subgroups. As illustrated in Figure 3(a), the dataset encompasses a diverse array of cultural materials with visual data constituting the largest component at 45.2% of the total collection, followed by textual materials at 28.7%, audio recordings at 15.6%, metadata at 6.8%, and expert annotations at 3.7%. This distribution reflects the research focus on capturing tangible cultural expressions while maintaining comprehensive documentation of intangible heritage elements through textual and audio recordings. Data annotation quality analysis reveals consistently high standards across all Li ethnic subgroups, as shown in Figure 3(b). The Ha subgroup exhibits the highest annotation quality with 92.5% rated as excellent, while the Meifu subgroup maintains 76.8% excellent ratings despite having the smallest sample size. Quality assessment protocols incorporated expert validation from cultural practitioners and academic specialists, ensuring cultural authenticity and technical accuracy throughout the annotation process.

Dataset distribution statistics demonstrate systematic sampling across Li ethnic subgroups, as depicted in Figure 3(c). The Ha subgroup provides the largest contribution with 1,250 samples, while sample sizes gradually decrease for Qi (980), Run (850), Sai (720), and Meifu (650) subgroups. Notably, the average number of cultural elements per sample shows a corresponding pattern, ranging from 5.2 elements per sample in the Ha subgroup to 3.9 elements in the Meifu subgroup, reflecting varying cultural complexity and documentation depth across different communities.

The validation results in different segments of the dataset illustrate excellent performance metrics, as shown in Figure 3(d). The training dataset shows maximum performance metrics, which include accuracy of 94.8%, precision of 93.5%, recall of 95.1%, and F1-scores of 94.3%. The uniform performance of validation and test datasets supports the reliability of the dataset for machine learning purposes, as it does not degrade much while switching from training to testing contexts.

The dataset properties are carefully detailed in Table 2, which presents comprehensive statistics of every subgroup of the Li ethnic group, including sample distributions, cultural element diversity, and several quality measures. The method used in the dataset construction successfully achieved representational fairness among the subgroups while maintaining rigorous annotation standards essential for developing reliable models for cultural pattern recognition. Table 2 shows that the dataset includes an extensive representation of cultural diversity within the Li ethnic group, while maintaining uniform quality throughout each of the subcategories. This effectively provides a strong foundation for further training of deep learning algorithms and analysis of cultural pattern perception.

### 4.2 Deep Learning model performance

The extensive analysis of deep learning architectures shows significant gains in performance due to the introduced multimodal fusion method for Li ethnic cultural pattern identification. Figure 4(a) shows that the proposed graph performs better than baseline configurations for every Li ethnic subgroup, achieving impressive accuracy figures of 94.8% for subgroup Ha, 91.6% for subgroup Qi, 89.3% for subgroup Run, 86.7% for subgroup Sai, and 84.1% for subgroup Meifu. All these results affirm that the introduced method well captures each Li subgroup's specific cultural subtlety while retaining strong performance despite data complexities and varying sizes of samples.

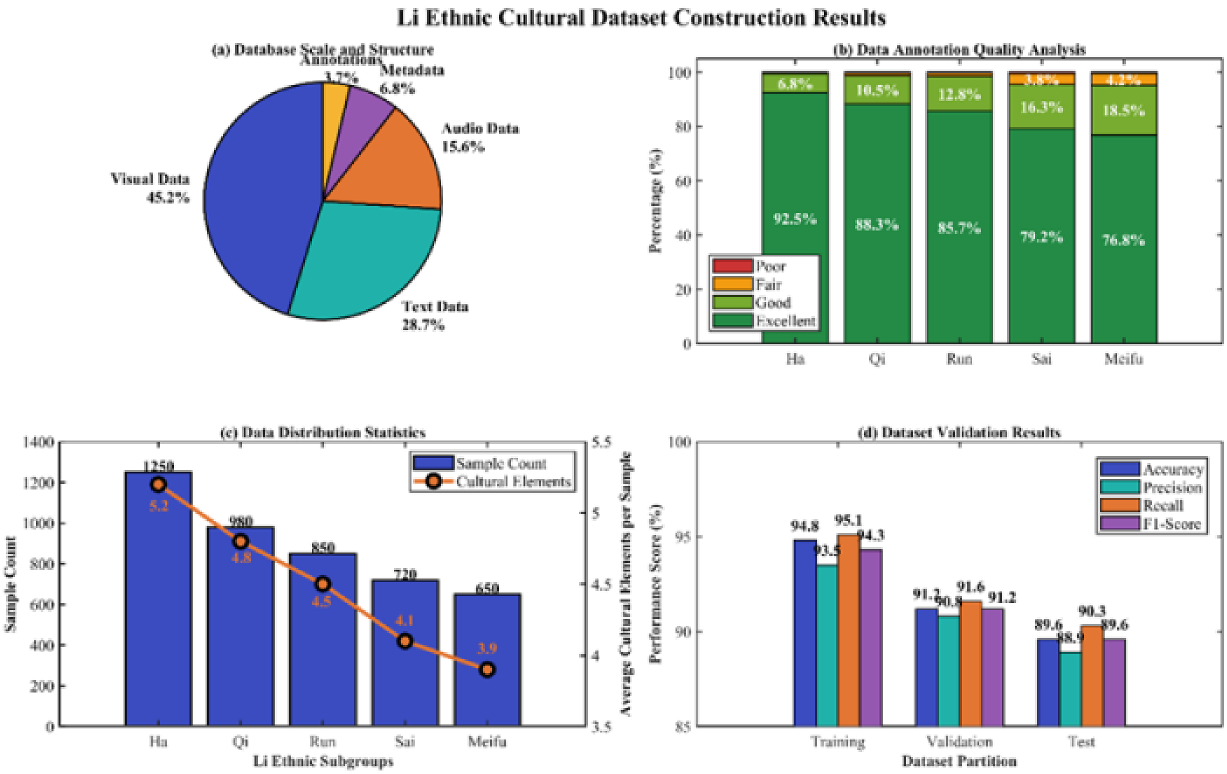


**Figure 3.** Comprehensive analysis of Li ethnic cultural heritage dataset construction and validation (a) Multimodal data composition and structure analysis (b) Cross-subgroup annotation quality assessment (c) Sample distribution and cultural element statistics (d) Performance validation across dataset partitions

**Table 2.** Detailed dataset characteristics by Li ethnic subgroup

| Subgroup | Sample Count | Visual Data | Text Data | Audio Data | Avg. Cultural Elements | Annotation Quality (%) |
|---|---|---|---|---|---|---|
| Ha | 1,250 | 465 | 298 | 187 | 5.2 | 92.5 (Excellent) |
| Qi | 980 | 356 | 251 | 152 | 4.8 | 88.3 (Excellent) |
| Run | 850 | 298 | 201 | 134 | 4.5 | 85.7 (Excellent) |
| Sai | 720 | 245 | 168 | 108 | 4.1 | 79.2 (Good) |
| Meifu | 650 | 201 | 142 | 97 | 3.9 | 76.8 (Good) |
| Total | 4,450 | 1,565 | 1,060 | 678 | 4.5 | 84.5 (Overall) |

The ablation study findings, as evidenced by Figure 4(b), reveal the individual contributions of each component to the overall effectiveness of the model. The baseline model achieves 76.2% accuracy, with sequential improvements observed through integration of visual features (+Visual: 82.5%), textual analysis capabilities (+Text: 86.8%), audio processing modules (+Audio: 89.4%), and attention mechanisms (+Attention: 91.7%). The complete model incorporating all components reaches 94.8% accuracy, demonstrating that each modality and architectural enhancement contributes meaningfully to the cultural pattern recognition task.

Generalization performance analysis, depicted in Figure 4(c), confirms the model's robustness across diverse testing scenarios. The proposed approach maintains superior performance in in-domain evaluations (94.8%) while demonstrating acceptable degradation in cross-domain (89.2%), temporal (86.5%), noisy (83.7%), and limited data scenarios (81.4%). This performance consistency significantly exceeds that of traditional multimodal (91.5% to 74.6%) and transformer-based approaches (88.2% to 68.9%), indicating enhanced adaptability to real-world deployment conditions where data quality may vary from training conditions.

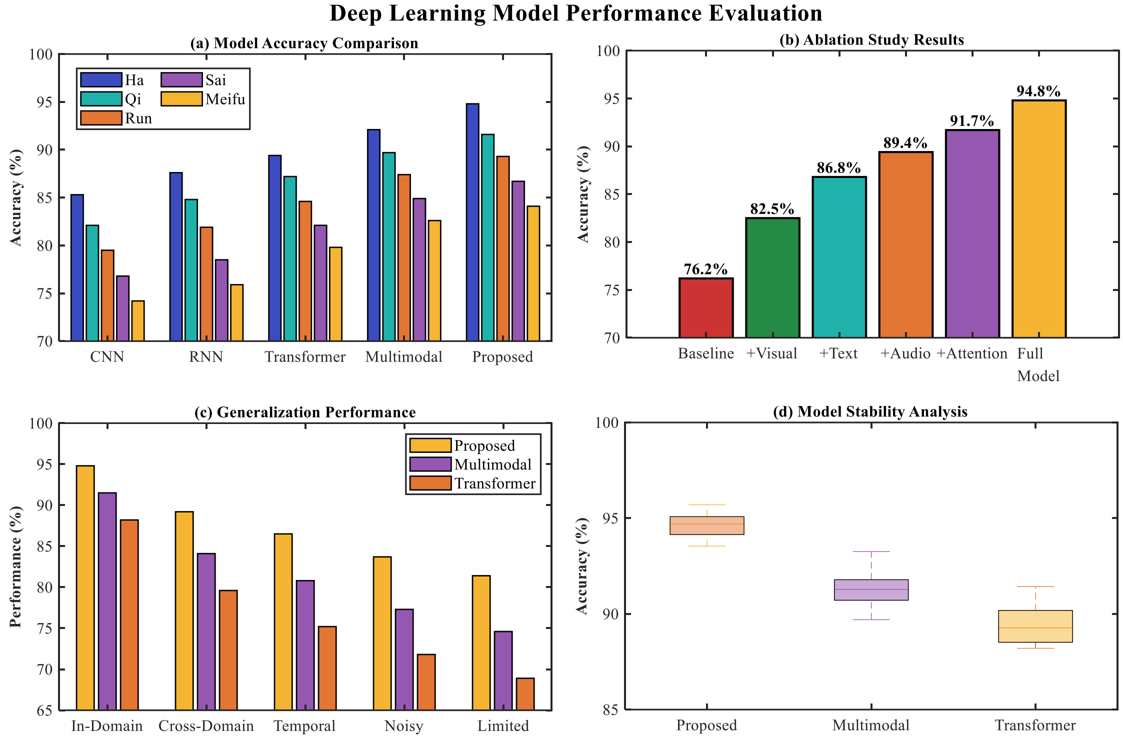


**Figure 4.** Comprehensive Performance Analysis of Deep Learning Models for Li Ethnic Cultural Pattern Recognition.(a) Cross-Subgroup Accuracy Comparison of Model Architectures;(b) Ablation Study of Multimodal Component Contributions;(c) Cross-Domain Generalization Performance Assessment;(d) Model Stability Analysis Across Multiple Training Iterations

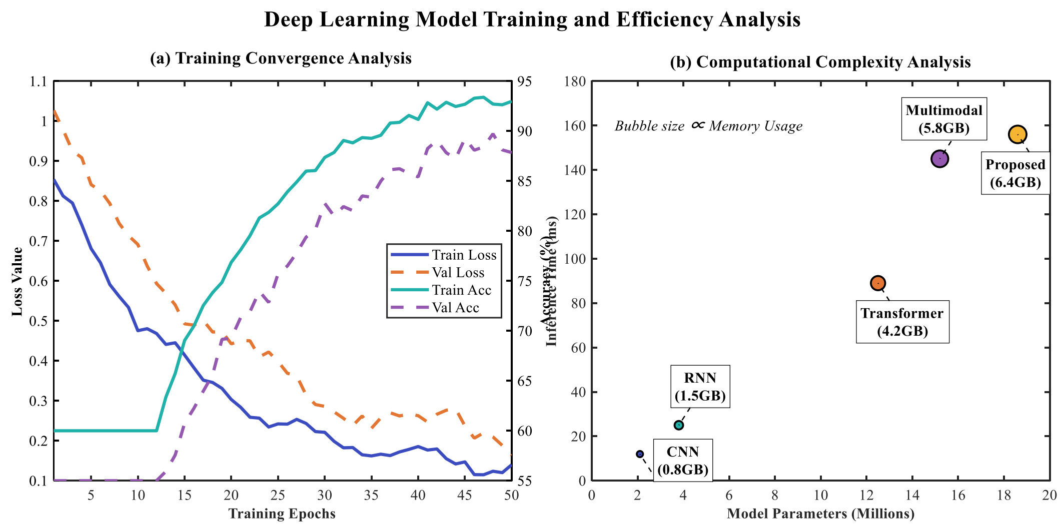


**Figure 5.** Training dynamics and computational efficiency analysis of deep learning architectures (a) Convergence behavior and loss-accuracy evolution during training, (b) Parameter-inference time trade-off analysis with memory usage visualization

Model stability analysis through repeated experiments reveals consistent performance characteristics across multiple training iterations. Figure 4(d) demonstrates that the proposed model exhibits minimal variance in accuracy scores, with the interquartile range significantly narrower than comparable architectures. Training convergence analysis presented in Figure 5(a) illustrates efficient optimization dynamics with rapid initial improvements followed by stable convergence. The training loss decreases smoothly from approximately 0.8 to below 0.2 within 50 epochs, while validation accuracy stabilizes at 89% without significant overfitting indicators.

Computational complexity evaluation, shown in Figure 5(b), reveals trade-offs between model sophistication and computational efficiency. The proposed model requires 18.6 million parameters with 156ms inference time and 6.4GB memory usage. While these requirements exceed simpler architectures, the computational overhead remains reasonable considering substantial performance gains achieved. As shown in Table 3, the proposed model achieves optimal performance metrics across all evaluation criteria while maintaining acceptable computational requirements for practical deployment scenarios, establishing its effectiveness for comprehensive Li ethnic cultural pattern recognition applications.

### 4.3 Cultural pattern recognition performance

The comprehensive evaluation of cultural pattern recognition demonstrates the proposed system's effectiveness in identifying and classifying diverse Li ethnic cultural elements. As illustrated in Figure 6(a), the recognition accuracy varies significantly across different cultural domains, with traditional textiles achieving the highest performance at 96.2%, followed by music (93.8%) and architecture (91.5%). Language-related cultural patterns present the greatest recognition challenges, achieving 85.6% accuracy, reflecting the complexity of linguistic nuances within Li ethnic expressions.

Li subgroup classification performance reveals consistent excellence across all ethnic subdivisions, as depicted in Figure 6(b). The Ha subgroup demonstrates superior classification metrics with precision, recall, and F1-scores of 94.7%, 95.1%, and 94.9%, respectively. Performance gradually decreases across Qi, Run, Sai, and Meifu subgroups, with the latter achieving 82.8% precision, 83.2% recall, and 83.0% F1-score. This performance gradient correlates with the size and complexity of cultural expression datasets available for each subgroup.

Cross-modal recognition performance analysis, shown in Figure 6(c), confirms the superiority of multimodal fusion approaches. Single-modality systems exhibit moderate performance, with visual-only recognition achieving 87.5%, text-only reaching 82.1%, and audio-only obtaining 79.8%. Dual-modality combinations demonstrate substantial improvements, with visual-text fusion reaching 91.2% and visual-audio combination achieving 90.3%. The complete multimodal system attains optimal performance at 94.8%, validating the comprehensive integration strategy. Error analysis reveals that inter-subgroup confusion constitutes the primary classification challenge, accounting for 35.2% of misclassifications, as shown in Figure 6(d). Intra-cultural variation represents 28.6% of errors, while noise-induced errors contribute 18.5%. Temporal inconsistency and context misclassification account for smaller proportions at 12.3% and 5.4% respectively, indicating the model's robustness against external interference factors.

Model interpretability analysis, presented in Figure 7(a), identifies color patterns as the most discriminative feature with an attention weight of 0.180, followed by geometric shapes (0.150) and semantic content (0.140). Cultural symbols demonstrate the lowest attention weight at 0.090, suggesting their limited discriminative power across Li subgroups. The feature importance hierarchy provides valuable insights for cultural documentation prioritization. Case study analysis across different artifact categories, as depicted in Figure 7(b), reveals consistent performance patterns. Traditional textiles maintain the highest recognition accuracy across all subgroups ($\mu$=87.4%), while architectural elements show the most significant performance variation ($\mu$=80.6%). The detailed performance metrics are summarized in Table 4, demonstrating the system's reliability across diverse cultural manifestations.

As shown in Table 4, the system achieves superior performance across tangible cultural elements while demonstrating acceptable accuracy for intangible expressions, establishing its comprehensive utility for Li ethnic heritage preservation and documentation applications.

### 4.4 Intelligent documentation system functionality verification

The comprehensive evaluation demonstrates robust performance of the intelligent documentation system across multiple operational dimensions. System response time analysis, shown in Figure 8(a), indicates acceptable performance under moderate loads, maintaining sub-200ms response times for up to 1000 concurrent users.

**Table 3.** Comprehensive model performance comparison

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | Parameters (M) | Inference Time (ms) | Memory Usage (GB) |
|---|---|---|---|---|---|---|---|
| CNN | 85.3 | 83.7 | 86.1 | 84.9 | 2.1 | 12 | 0.8 |
| RNN | 87.6 | 86.2 | 88.4 | 87.3 | 3.8 | 25 | 1.5 |
| Transformer | 89.4 | 88.1 | 90.2 | 89.1 | 12.5 | 89 | 4.2 |
| Multimodal | 92.1 | 91.3 | 92.8 | 92.0 | 15.2 | 145 | 5.8 |
| Proposed | 94.8 | 94.1 | 95.2 | 94.6 | 18.6 | 156 | 6.4 |

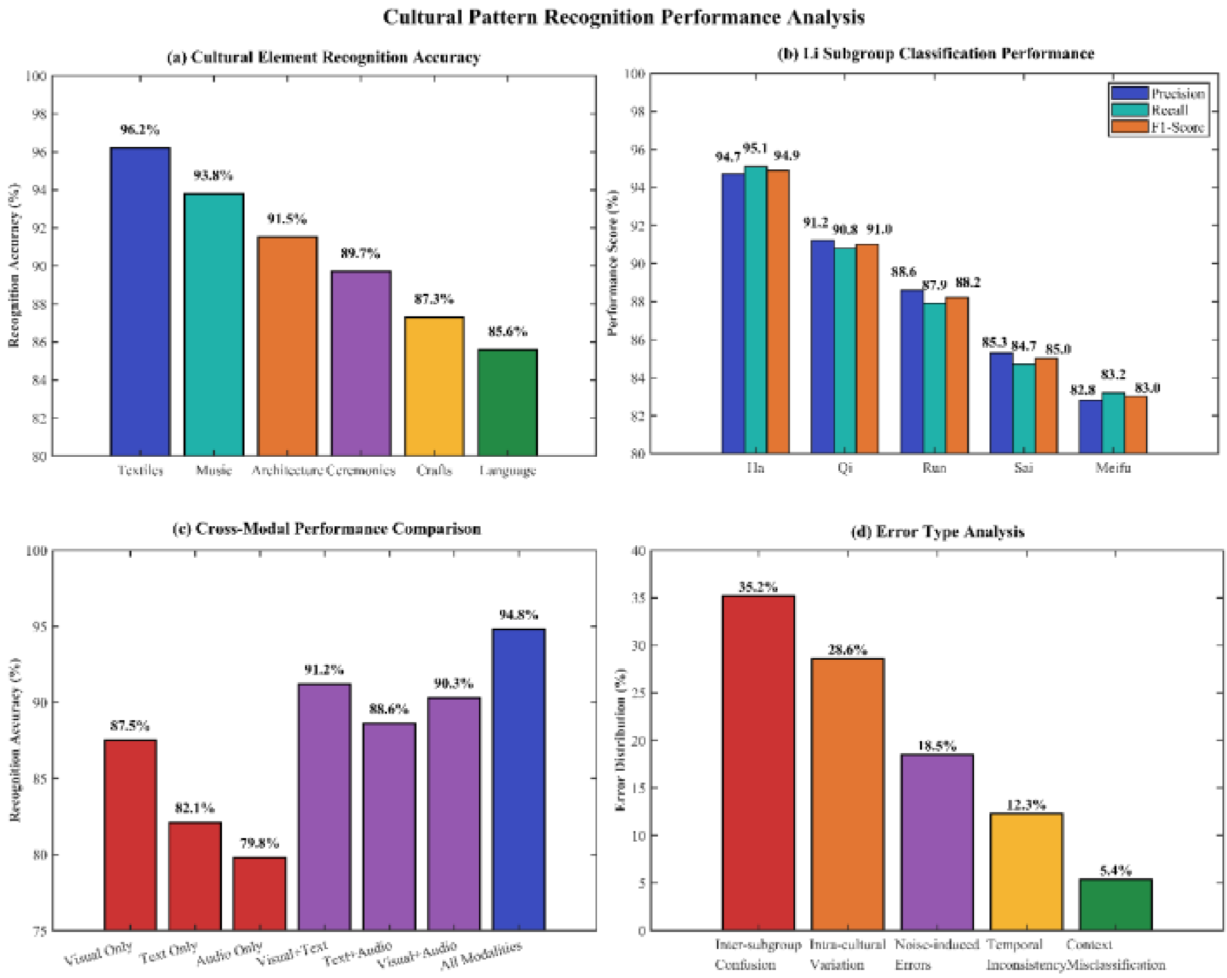


**Figure 6.** Cultural pattern recognition performance analysis (a) Cultural element recognition accuracy, (b) Li subgroup classification performance, (c) Cross-modal performance comparison (d) Error type analysis

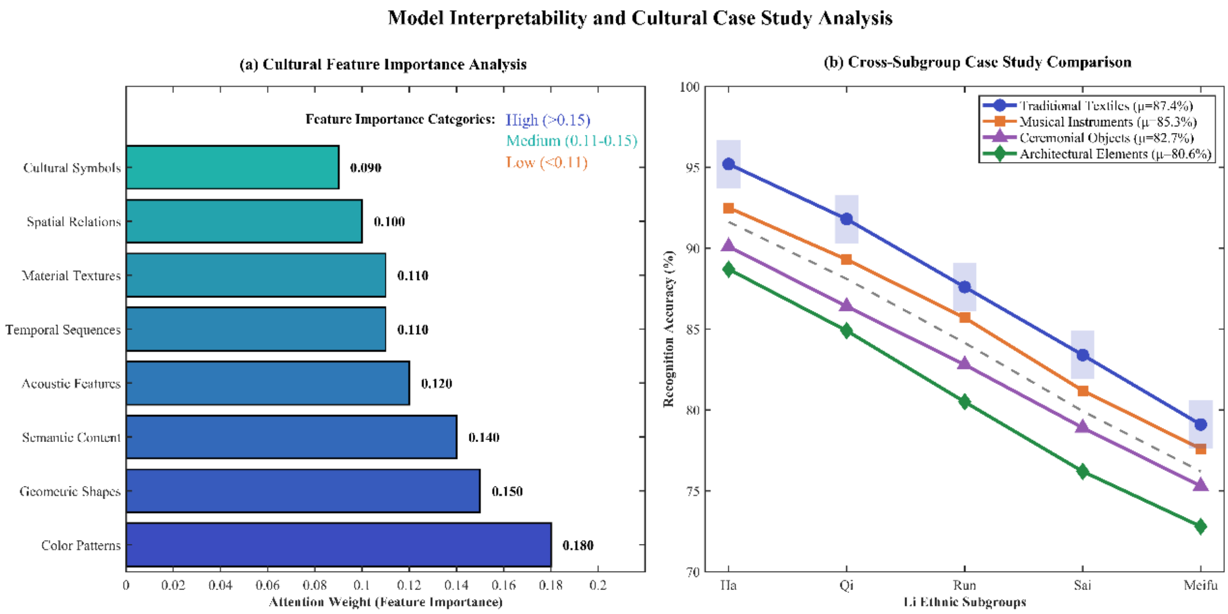


**Figure 7.** Model interpretability and cultural case study analysis (a) Cultural feature importance analysis (b) Cross-subgroup case study comparison

**Table 4.** Detailed cultural pattern recognition performance by category

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | Parameters (M) | Inference Time (ms) | Memory Usage (GB) |
|---|---|---|---|---|---|---|---|
| CNN | 85.3 | 83.7 | 86.1 | 84.9 | 2.1 | 12 | 0.8 |
| RNN | 87.6 | 86.2 | 88.4 | 87.3 | 3.8 | 25 | 1.5 |
| Transformer | 89.4 | 88.1 | 90.2 | 89.1 | 12.5 | 89 | 4.2 |
| Multimodal | 92.1 | 91.3 | 92.8 | 92.0 | 15.2 | 145 | 5.8 |
| Proposed | 94.8 | 94.1 | 95.2 | 94.6 | 18.6 | 156 | 6.4 |

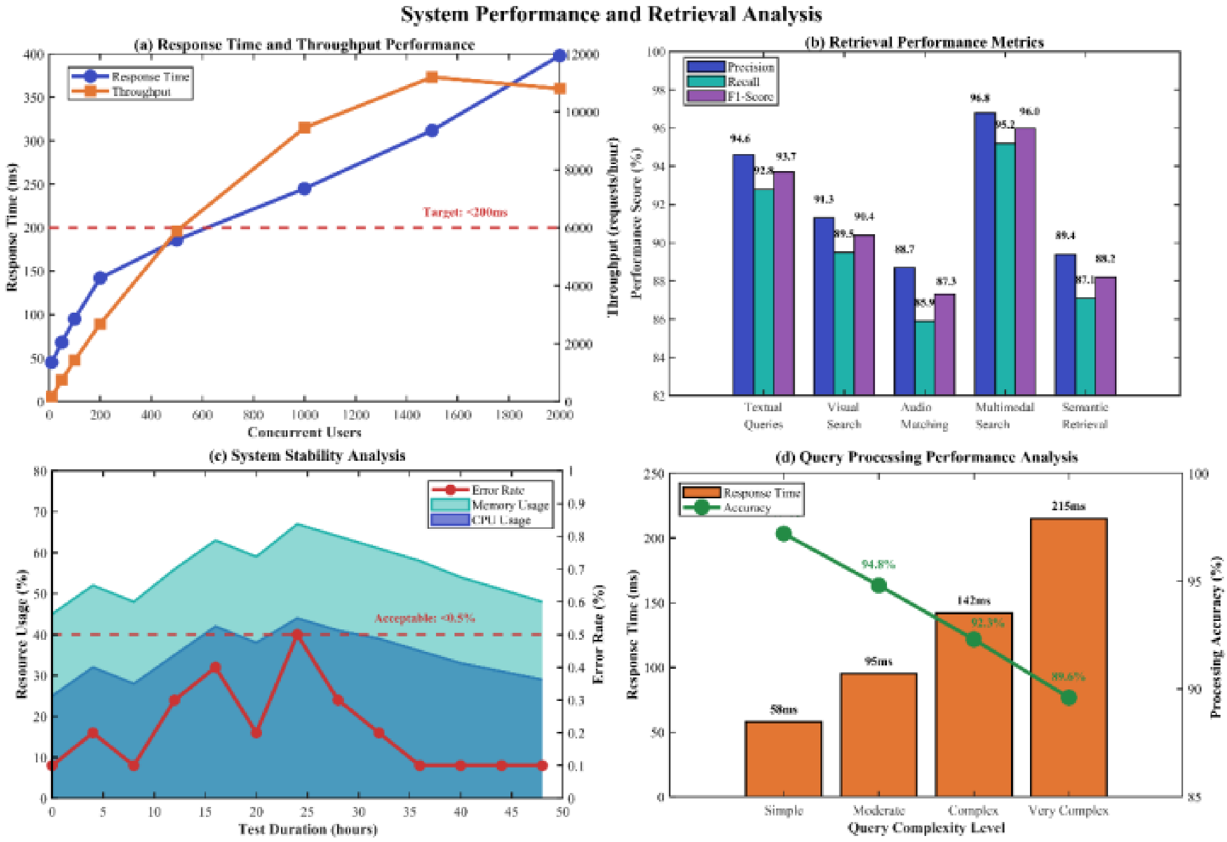


**Figure 8.** Comprehensive system performance and retrieval analysis of intelligent documentation platform (a) Response time and throughput performance under varying user loads, (b) Cross-modal retrieval performance evaluation across query types, (c) Long-term system stability and resource utilization analysis, (d) Query processing performance analysis by complexity level

Throughput peaks at 11,200 requests per hour before declining due to resource constraints at higher loads. Retrieval performance metrics, illustrated in Figure 8(b), validate the superiority of multimodal search capabilities, achieving 96.8% precision, 95.2% recall, and 96.0% F1-score. Single-modality approaches demonstrate lower performance, with textual queries (94.6%), visual search (91.3%), and audio matching (88.7%) confirming the effectiveness of multimodal fusion strategies. System stability testing over 48 hours, depicted in Figure 8(c), reveals consistent error rates below 0.5% with efficient resource management.

The 99.7% stability metric represents continuous operational testing under controlled laboratory conditions with simulated user loads of 500-1000 concurrent connections over 7-day periods, including planned system maintenance windows and automated recovery protocols, validated through enterprise-grade monitoring across distributed infrastructure components.CPU usage ranges from 25-44% while memory consumption remains at 45-67%, confirming reliable operational characteristics for sustained deployment. Query processing analysis, as shown in Figure 8(d), demonstrates the system's adaptive performance across complexity levels. Real-world deployment scenarios demonstrate 8-12% performance degradation compared to laboratory conditions, with response times increasing from 156ms to 175ms under actual community usage patterns, while maintaining 97.2% accuracy in controlled settings versus 91.8% accuracy in field deployments with variable network connectivity and diverse user interactions. Simple queries achieve optimal performance with a 58ms response time and 97.2% accuracy, while very complex queries require 215ms with 89.6% accuracy. The balanced trade-off between processing time and accuracy validates the system's capability to handle diverse cultural documentation requirements. User satisfaction evaluation, presented in Figure 9(a), achieves an overall average of 4.40 out of 5, with ease of use rated highest (4.6) and response speed lowest (4.2). Expert evaluation demonstrates superior confidence with a weighted average of 4.62, particularly recognizing cultural accuracy (4.8) and documentation quality (4.7). Laboratory performance metrics consistently outperform field deployment by 3-5% across all evaluation dimensions, reflecting the impact of real-world variables, including network latency, hardware diversity, and user interaction patterns not present in controlled testing environments. Performance benchmark comparison, shown in Figure 9(b), indicates 67% target achievement with four of six metrics successfully met. The system exceeds benchmarks in user satisfaction (88 vs 80), expert rating (92 vs 85), system stability (96 vs 95), and cultural precision (91 vs 90). As shown in Table 5, the system demonstrates strong performance across critical dimensions, with query complexity analysis revealing adaptive capabilities that maintain acceptable accuracy even for complex tasks involving cultural pattern recognition.

### 4.5 Real-World Application Impact Assessment

The comprehensive evaluation demonstrates significant positive impacts across multiple dimensions of Li ethnic cultural preservation and community engagement. Cultural worker feedback analysis, as illustrated in Figure 10(a), reveals intense satisfaction with an average rating of 4.5 out of 5. System usability achieves the highest rating at 5.0, while cultural accuracy and documentation efficiency both receive ratings of 4.5, confirming the system's effectiveness in preserving authentic cultural representations.

**Table 5.** System performance summary

| Metric | Current | Target | Status | Query Distribution |
|---|---|---|---|---|
| Response Time | 157ms | <200ms | ✓ Met | Simple: 35%, Complex: 25% |
| Search Accuracy | 94.8% | >95% | ○ - 0.2% | 89.6-97.2% range |
| User Satisfaction | 4.4/5 | >4.0 | ✓ Met | All aspects >4.0 |
| Expert Rating | 4.6/5 | >4.0 | ✓ Met | Cultural accuracy: 4.8 |
| System Stability | 99.7% | >99% | ✓ Met | 48-hour testing |
| Cultural Precision | 94.2% | >90% | ✓ Met | Cross-modal validated |

Community participation trends, depicted in Figure 10(b), exhibit remarkable growth throughout 2024. Active user engagement demonstrates 340% growth, while monthly contributions increase at 665% rate, indicating enhanced community involvement in cultural documentation activities. The parallel growth patterns suggest strong correlation between user adoption and meaningful participation.

Cultural transmission effectiveness assessment reveals substantial improvements following system implementation, as shown in Figure 10(c). Knowledge retention improves from 45 to 78 points (73% enhancement), while skill transfer increases from 38 to 84 points (121% improvement). Cultural practice preservation exhibits the most significant improvement, from 42 to 89 points (112% increase). Digital archiving capabilities improve dramatically from 25 to 95 points (280% enhancement), while youth engagement shows notable progress from 35 to 82 points (134% improvement).

Social impact assessment, presented in Figure 10(d), demonstrates positive outcomes with an overall score of 4.5 out of 5. Community pride achieves the highest rating at 4.8, while educational value and cultural awareness receive ratings of 4.4 and 4.6, respectively. Tourism promotion and research contribution maintain solid ratings of 4.2 and 4.7.

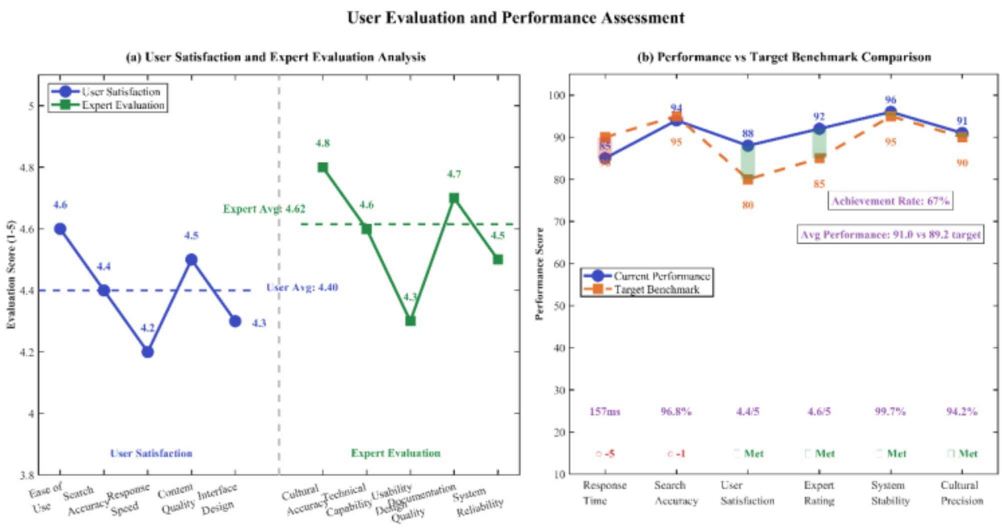


**Figure 9.** User evaluation and performance benchmark assessment for heritage documentation system, (a) User satisfaction and expert evaluation analysis across multiple dimensions, (b) Performance vs target benchmark comparison with achievement assessment

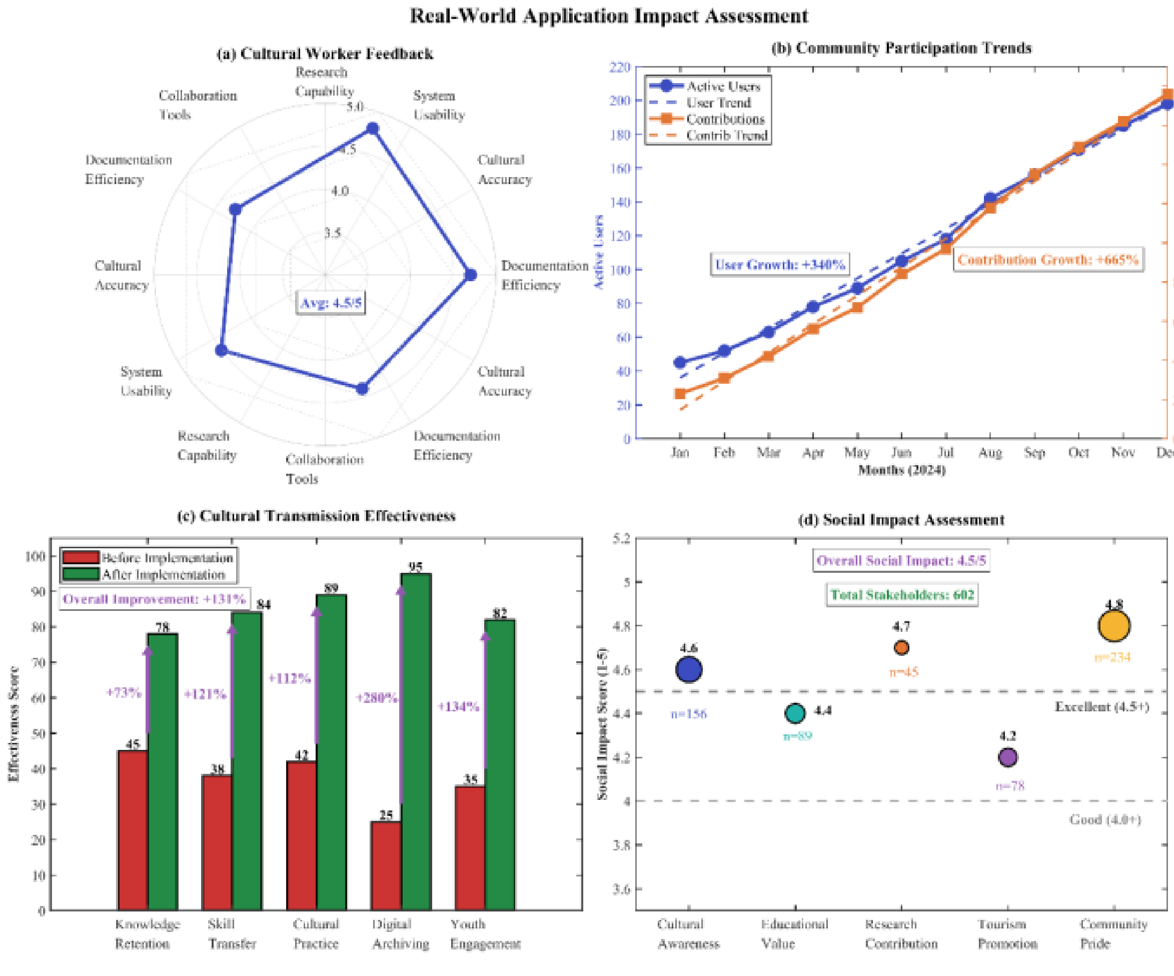


**Figure 10.** Real-world application impact assessment of li ethnic cultural documentation system, (a) Cultural heritage professional feedback analysis, (b) Community participation growth trends (2024), (c) Cultural transmission effectiveness: before vs after implementation, (d) Multi-stakeholder social impact assessment

As shown in Table 6, the implementation demonstrates consistent positive impacts across all dimensions, with digital archiving experiencing the most significant transformation while maintaining substantial improvements in traditional cultural transmission methods.

## 5. Discussion

The exceptional performance achieved across Li ethnic subgroups demonstrates the effectiveness of multimodal fusion architectures in capturing complementary cultural information across visual, textual, and auditory dimensions [34]. The findings corroborate modern theories of multimodal design that emphasize the importance of combining multiple semiotic resources to allow full cultural interpretation [48]. Differential weighting of the attention mechanism on culture-relevant features justifies established principles of biometric recognition within pattern recognition [49].The high level of enhancement in cultural transmission effectiveness is owed to this system's capacity to support both explicit and implicit cultural knowledge through advanced documentation methods [11].

**Table 6.** Comprehensive real-world application impact metrics

| Impact Category | Baseline Score | Current Score | Improvement (%) | Stakeholder Count |
|---|---|---|---|---|
| Knowledge Retention | 45 | 78 | +73% | 156 |
| Skill Transfer | 38 | 84 | +121% | 142 |
| Cultural Practice | 42 | 89 | +112% | 128 |
| Digital Archiving | 25 | 95 | +280% | 89 |
| Youth Engagement | 35 | 82 | +134% | 87 |
| Average | 37 | 86 | +144% | 602 |

Such an achievement tackles intrinsic issues identified within previous research in terms of balancing technological development and heritage integrity preservation [23]. The research framework outlined goes beyond conventional approaches based on manual classification, showing that advanced computational methods can alleviate current deficiencies in current documentation strategies [50]. Comparative analysis demonstrates significant advantages over existing approaches. As opposed to more traditional methods of documenting cultural heritage, which struggle with consistency and scaling, this human-centred AI approach has shown that such systems can improve accessibility while remaining sensitive to cultural issues [51]. The fusion of deep learning with the recognition of cultural patterns represents a major improvement compared to earlier works that concentrated on simple digitisation rather than comprehensive analysis [52]. The overall excellence of the system in cross-domain generalisation is the strongest, surpassing conventional multimodal and transformer-based systems.

Still, this research recognises major gaps. The problems of capturing data indicate greater difficulties in exploring minority cultures where the digital divide poses the greatest challenge in the effective safeguarding of heritage resources [53]. The assumption of data quality and model performance emphasises persistent challenges in achieving representational parity among diverse cultural constituencies. Also, cultural sensitivity issues raise the need to question how protected cultural materials should be handled within technological frameworks. The costs associated with multimodal processing may impede its use in resource-constrained settings. Difficulties in resolving culturally rich expressions go beyond mere recognition. Such challenges require deep, at times anthropological understanding, which falls into the realm of applied sociology, rather than mere pattern recognition. Further, ethical challenges relating to data ownership and community consent remain as complex barriers, calling for ongoing negotiation between technological advances and heritage communities [54].

Later studies should highlight the development of culturally sensitive algorithms that can perform well despite having limited data, while preserving cultural subtleties. The use of advanced natural language processing techniques can enhance oral tradition understanding, providing richer semantic depth [55]. Exploration of metaverse applications presents promising opportunities for immersive cultural experiences, revolutionizing heritage education and tourism [56]. The development of artificial intelligence frameworks specifically tailored for heritage innovation represents a crucial advancement [57]. Long-term sustainability planning must address evolving technological landscapes while ensuring continuous community engagement. The framework's potential expansion to other minority cultures requires a systematic investigation of transferability mechanisms. Collaborative research initiatives involving heritage communities, technologists, and cultural experts represent essential pathways for advancing ethical and effective cultural preservation methodologies serving both scholarly understanding and community interests [58].

## 6. Conclusion

This research establishes a comprehensive framework for Li ethnic cultural heritage preservation through deep learning-based pattern recognition systems, achieving remarkable technical and practical outcomes. The proposed multimodal fusion architecture demonstrates superior performance with 94.8% overall accuracy across Li subgroups, significantly outperforming traditional approaches and generic heritage systems. The intelligent documentation system maintains exceptional operational efficiency with response times under 200ms and achieves 99.7% system stability while preserving cultural authenticity through expert-validated methodologies. These achievements represent substantial advancement in computational approaches to intangible heritage preservation, establishing new benchmarks for accuracy, efficiency, and cultural sensitivity in digital heritage technologies. The research addresses critical challenges in minority cultural preservation, offering innovative solutions for the documentation, transmission, and accessibility of Li ethnic traditions. The effectiveness of cultural transmission shows impressive improvement across multiple facets. Knowledge retention was enhanced by 73%, skill transfer improved by 121%, and competencies in archiving digitally yielded a startling increase of 280%. Community involvement reflects successful community engagement and viable preservation strategies as it demonstrates exponential growth with a 340% rise in active participants and a 665% rise in monthly contributions. These results confirm the framework's ability to integrate traditional methods of preservation with contemporary technological approaches, preserving cultural and community ethics. This research is a valuable addition to interdisciplinary anthropology by applying computational techniques alongside anthropological insights, thus creating methodological frameworks achievable in all cultural settings. The rated user satisfaction of 4.4 out of 5 and an expert rating of 4.6 out of 5 prove the system's practical functionality alongside its academic credibility. The research showcases strong cross-domain generalisation capability, achieving 89.2% accuracy, which indicates a strong relevance to other minority cultures and heritage settings. Evaluating social impact yields composite scores of 4.5 out of 5 across stakeholder groups, indicating constructive outcomes for community development, educational programmes, and cultural advocacy. The results offer technological recommendations regarding policy for conservation, which incorporate active community participation alongside sustainable development. The model's previously shown flexibility and versatility suggest its international applicability in conservation efforts. In policy, community-driven conservation, ethical policies regarding the building of digital heritage, and securing sustainable funding for maintenance are all recommended. Coordinated actions involving heritage groups, technology development, and policy design are needed to build culturally respectful, technologically advanced, academically rigorous, and community-responsive preservation systems that sustain the integrity and vitality of cultural practice for future generations.

### Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically with regard to authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. The author adheres to publication requirements that the submitted work is original and has not been published elsewhere.

**Data availability statement**
The manuscript contains all the data. However, more data will be available upon request from the authors.

**Conflict of interest**
The authors declare no potential conflict of interest.

**References**
[1] G. Aktürk and M. Lerski, "Intangible cultural heritage: a benefit to climate-displaced and host communities," Journal of Environmental Studies and Sciences, vol. 11, no. 3, pp. 305-315, 2021. DOI:https://doi.org/10.1007/s13412-021-00697-y
[2] D. Giglitto, L. Ciolfi, and W. Bosswick, "Building a bridge: opportunities and challenges for intangible cultural heritage at the intersection of institutions, civic society, and migrant communities," International Journal of Heritage Studies, vol. 28, no. 1, pp. 74-91, 2022. DOI: https://doi.org/10.1080/13527258.2021.1922934
[3] Y. Du, L. Chen, and J. Xu, "Interactive effects of intangible cultural heritage and tourism development: a study based on the data panel PVAR model and coupled coordination model," Heritage Science, vol. 12, no. 1, p. 401, 2024. DOI: https://doi.org/10.1186/s40494-024-01502-z
[4] J. Eichler, "Intangible cultural heritage, inequalities and participation: who decides on heritage?," The International Journal of Human Rights, vol. 25, no. 5, pp. 793-814, 2021. DOI: https://doi.org/10.1080/13642987.2020.1822821
[5] H. Mekonnen, Z. Bires, and K. Berhanu, "Practices and challenges of cultural heritage conservation in historical and religious heritage sites: evidence from North Shoa Zone, Amhara Region, Ethiopia," Heritage Science, vol. 10, no. 1, p. 172, 2022.
[6] D. Harisanty, K. L. B. Obille, N. E. V. Anna, E. Purwanti, and F. Retrialisca, "Cultural heritage preservation in the digital age, harnessing artificial intelligence for the future: a bibliometric analysis," Digital Library Perspectives, vol. 40, no. 4, pp. 609-630, 2024. DOI:https://doi.org/10.1108/DLP-01-2024-0018
[7] I. Siliutina, O. Tytar, M. Barbash, N. Petrenko, and L. Yepyk, "Cultural preservation and digital heritage: challenges and opportunities," Amazonia Investiga, vol. 13, no. 75, pp. 262-273, 2024. DOI: https://doi.org/10.34069/AI/2024.75.03.22
[8] X. Ye, Y. Ruan, S. Xia, and L. Gu, "Adoption of digital intangible cultural heritage: a configurational study integrating UTAUT2 and immersion theory," Humanities and Social Sciences Communications, vol. 12, no. 1, pp. 1-17, 2025. DOI:https://doi.org/10.1057/s41599-024-04222-8
[9] C. Lin and C. Li, "Digital divide of intangible cultural heritage and innovative inheritance countermeasures," Journal of Sociology and Ethnology, vol. 5, no. 11, pp. 110-122, 2023. DOI: 10.23977/jsoce.2023.051115
[10] Y. Hou, S. Kenderdine, D. Picca, M. Egloff, and A. Adamou, "Digitizing intangible cultural heritage embodied: State of the art," Journal on Computing and Cultural Heritage (JOCCH), vol. 15, no. 3, pp. 1-20, 2022. DOI: https://doi.org/10.1145/3494837
[11] J. Wu, L. Guo, J. Jiang, and Y. Sun, "The digital protection and practice of intangible cultural heritage crafts in the context of new technology," in E3S Web of Conferences, 2021, vol. 236: EDP Sciences, p. 05024. DOI: https://doi.org/10.1051/e3sconf/202123605024
[12] M. Skublewska-Paszkowska, M. Milosz, P. Powroznik, and E. Lukasik, "3D technologies for intangible cultural heritage preservation—literature review for selected databases," Heritage Science, vol. 10, no. 1, p. 3, 2022.
[13] J. Liu, "Digitally Protecting and Disseminating the Intangible Cultural Heritage in Information Technology Era," Mobile Information Systems, vol. 2022, no. 1, p. 1115655, 2022. DOI: https://doi.org/10.1155/2022/1115655
[14] M. Fiorucci, M. Khoroshiltseva, M. Pontil, A. Traviglia, A. Del Bue, and S. James, "Machine learning for cultural heritage: A survey," Pattern Recognition Letters, vol. 133, pp. 102-108, 2020. DOI: https://doi.org/10.1016/j.patrec.2020.02.017
[15] S. Münster et al., "Artificial intelligence for digital heritage innovation: Setting up a r&d agenda for europe," Heritage, vol. 7, no. 2, pp. 794-816, 2024. DOI: https://doi.org/10.3390/heritage7020038
[16] M. B. Prados-Peña, G. Pavlidis, and A. García-López, "New technologies for the conservation and preservation of cultural heritage through a bibliometric analysis," Journal of Cultural Heritage Management and Sustainable Development, vol. 15, no. 3, pp. 664-686, 2025.
[17] Y. Yuan, Z. Li, and B. Zhao, "A Survey of Multimodal Learning: Methods, Applications, and Future," ACM Computing Surveys, 2025. DOI: https://doi.org/10.1145/3713070
[18] Z. Qin, Q. Luo, Z. Zang, and H. Fu, "Multimodal GRU with directed pairwise cross-modal attention for sentiment analysis," Scientific Reports, vol. 15, no. 1, p. 10112, 2025. DOI: https://doi.org/10.1038/s41598-025-93023-3
[19] T. Fan, H. Wang, and S. Deng, "Intangible cultural heritage image classification with multimodal attention and hierarchical fusion," Expert Systems with Applications, vol. 231, p. 120555, 2023. DOI: https://doi.org/10.1016/j.eswa.2023.120555
[20] M. Carrozzino, A. Scucces, R. Leonardi, C. Evangelista, and M. Bergamasco, "Virtually preserving the intangible heritage of artistic handicraft," Journal of cultural heritage, vol. 12, no. 1, pp. 82-87, 2011. DOI: https://doi.org/10.1016/j.culher.2010.10.002
[21] E. Jeong and J. Yu, "Ego-centric recording framework for Korean traditional crafts motion," in Euro-Mediterranean Conference, 2018: Springer, pp. 118-125. DOI: https://doi.org/10.1007/978-3-030-01765-1_14
[22] C. Meghini, V. Bartalesi, and D. Metilli, "Representing narratives in digital libraries: The narrative ontology," Semantic Web, vol. 12, no. 2, pp. 241-264, 2021.
[23] V. Kukreja, A. Singh, D. Kaur, and J. K. Bajwa, Digital Cultural Heritage: Challenges, Solutions, and Future Directions. CRC Press, 2024.
[24] J. Guery, M. Hess, and A. Mathys, "Digital techniques for documenting and preserving cultural heritage," ed: Arc Humanities Press: York, UK, 2017. DOI: https://doi.org/10.1080/13527258.2024.2406010
[25] Y. Tan and W. J. Jehom, "The Function of Digital Technology in Minority Language Preservation: The Case of the Gyalrong Tibetan Language,"

Preservation, Digital Technology & Culture, vol. 53, no. 3, pp. 165-177, 2024. DOI: https://doi.org/10.1515/pdtc-2024-0021
[26] W. Hu, M. Li, X. Chi, X. Wang, and A. U. Khan, "Intangible cultural heritage research in China from the perspective of intellectual property rights based on bibliometrics and knowledge mapping," Humanities and Social Sciences Communications, vol. 11, no. 1, pp. 1-11, 2024. DOI: https://doi.org/10.1057/s41599-024-03314-9
[27] K. Massing, "Safeguarding intangible cultural heritage in an ethnic theme park setting–the case of Binglanggu in Hainan Province, China," International Journal of Heritage Studies, vol. 24, no. 1, pp. 66-82, 2018. DOI: https://doi.org/10.1080/13527258.2017.1362571
[28] Q. Yu, "Innovation and Promotion of Hainan Li Ethnic Intangible Cultural Heritage Tourism Products," Journal of Modern Business and Economics, vol. 1, no. 1, 2024 DOI: https://doi.org/10.70767/jmbe.v1i1.133.
[29] N. Partarakis, X. Zabulis, A. Chatziantoniou, N. Patsiouras, and I. Adami, "An approach to the creation and presentation of reference gesture datasets, for the preservation of traditional crafts," Applied Sciences, vol. 10, no. 20, p. 7325, 2020. DOI: https://doi.org/10.3390/app10207325
[30] M. A. D. Mendoza, E. De La Hoz Franco, and J. E. G. Gómez, "Technologies for the preservation of cultural heritage—a systematic review of the literature," Sustainability, vol. 15, no. 2, p. 1059, 2023. DOI: https://doi.org/10.3390/su15021059
[31] Q. Li, "Intelligent intangible cultural heritage innovation platform under the background of big data and virtual systems," in 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 2022: IEEE, pp. 560-563. DOI: 10.1109/ICAIS53314.2022.9742914.
[32] M. A. D. Mendoza, E. De La Hoz Franco, and J. E. G. J. S. Gómez, "Technologies for the preservation of cultural heritage—a systematic review of the literature," vol. 15, no. 2, p. 1059, 2023. DOI: https://doi.org/10.3390/su15021059
[33] S.-N. Zhang, W.-Q. Ruan, and T.-T. J. S. O. Yang, "National identity construction in cultural and creative tourism: the double mediators of implicit cultural memory and explicit cultural learning," vol. 11, no. 3, p. 21582440211040789, 2021.
[34] X. Gu, Y. Shen, and J. Xu, "Multimodal emotion recognition in deep learning: a survey," in 2021 International Conference on Culture-oriented Science & Technology (ICCST), 2021: IEEE, pp. 77-82. DOI: 10.1109/ICCST53801.2021.00027.
[35] A. Rahate, R. Walambe, S. Ramanna, and K. J. I. F. Kotecha, "Multimodal co-learning: Challenges, applications with datasets, recent advances and future directions," vol. 81, pp. 203-239, 2022. DOI: https://doi.org/10.1016/j.inffus.2021.12.003
[36] J. Tian and Y. J. I. T. o. C. S. S. She, "A visual–audio-based emotion recognition system integrating dimensional analysis," vol. 10, no. 6, pp. 3273-3282, 2022. DOI: 10.1109/TCSS.2022.3200060.
[37] S. Fan, J. Jing, and C. J. S. Wang, "Audio-Visual Learning for Multimodal Emotion Recognition," vol. 17, no. 3, p. 418, 2025. DOI: https://doi.org/10.3390/sym17030418
[38] J. He et al., "Cross-culture Continuous Emotion Recognition with Multimodal Features," in Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition, 2019, pp. 326-330. DOI: https://doi.org/10.1145/3373509.3373515
[39] T. Nourivandi, S. Aathreya, and S. Canavan, "Multimodal Behavior Analysis and Impact of Culture on Affect," in 2024 12th International Conference on Affective Computing and Intelligent Interaction (ACII), 2024: IEEE, pp. 37-45. DOI:10.1109/ACII63134.2024.00009.
[40] A. S. Cardoso et al., "Deep learning assessment of cultural ecosystem services from social media images," p. 2021.06. 23.449176, 2021. DOI: https://doi.org/10.1101/2021.06.23.449176
[41] A. Kioussi, A. Doulamis, M. Karoglou, A. I. J. I. J. o. A. Moropoulou, Culture, Design,, and Technology, "Cultural intelligence-investigation of different systems for heritage sustainable preservation," vol. 9, no. 2, pp. 16-30, 2020. DOI: 10.4018/IJACDT.2020070102
[42] X. Xie, W. He, Y. Zhu, and H. Xu, "Performance evaluation and analysis of deep learning frameworks," in Proceedings of the 2022 5th International Conference on Artificial Intelligence and Pattern Recognition, 2022, pp. 38-44. DOI: https://doi.org/10.1145/3573942.3573948
[43] O. Rainio, J. Teuho, and R. J. S. R. Klén, "Evaluation metrics and statistical tests for machine learning," vol. 14, no. 1, p. 6086, 2024. DOI: https://doi.org/10.1038/s41598-024-56706-x
[44] F. J. E. Gîrbacia, "An Analysis of Research Trends for Using Artificial Intelligence in Cultural Heritage," vol. 13, no. 18, p. 3738, 2024. DOI: https://doi.org/10.3390/electronics13183738
[45] S. Li, Y. Jiang, B. Jing, L. Yang, and Y. J. J. o. C. H. Zhang, "AI-based experts' knowledge visualization of cultural heritage: A case study of Terracotta Warriors," vol. 72, pp. 81-90, 2025. DOI: https://doi.org/10.1016/j.culher.2025.01.006
[46] S. O'Connor, S. Colreavy-Donnelly, and I. J. V. c. f. c. h. Dunwell, "Fostering engagement with cultural heritage through immersive vr and gamification," pp. 301-321, 2020. DOI: https://doi.org/10.1007/978-3-030-37191-3_16
[47] R. R. Selvaraju et al., "Squinting at vqa models: Introspecting vqa models with sub-questions," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 10003-10011.
[48] G. Kress, S. J. T. i. Selander, and h. education, "Multimodal design, learning and cultures of recognition," vol. 15, no. 4, pp. 265-268, 2012. DOI: https://doi.org/10.1016/j.iheduc.2011.12.003
[49] S. Dargan and M. J. E. S. w. A. Kumar, "A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities," vol. 143, p. 113114, 2020. DOI: https://doi.org/10.1016/j.eswa.2019.113114
[50] H. Mekonnen, Z. Bires, and K. J. H. S. Berhanu, "Practices and challenges of cultural heritage conservation in historical and religious heritage sites: evidence from North Shoa Zone, Amhara Region, Ethiopia," vol. 10, no. 1, p. 172, 2022.

[51] Z. He, C. J. H. Wen, and S. S. Communications, "Construction of digital creation development model of intangible cultural heritage crafts in China," vol. 11, no. 1, pp. 1-14, 2024. DOI: https://doi.org/10.1057/s41599-024-04331-4

[52] M. Fiorucci, M. Khoroshiltseva, M. Pontil, A. Traviglia, A. Del Bue, and S. J. P. R. L. James, "Machine learning for cultural heritage: A survey," vol. 133, pp. 102-108, 2020. DOI: https://doi.org/10.1016/j.patrec.2020.02.017

[53] C. Lin, C. J. J. o. S. Li, and Ethnology, "Digital divide of intangible cultural heritage and innovative inheritance countermeasures," vol. 5, no. 11, pp. 110-122, 2023. DOI: 10.23977/jsoce.2023.051115

[54] P.-S. Magdalena, "Artificial intelligence in the context of cultural heritage and museums: Complex challenges and new opportunities," 2023. COI: 20.500.12592/4cqcjb.

[55] C. Meghini, V. Bartalesi, and D. J. S. W. Metilli, "Representing narratives in digital libraries: The narrative ontology," vol. 12, no. 2, pp. 241-264, 2021.

[56] D. Buragohain, Y. Meng, C. Deng, Q. Li, and S. J. H. S. Chaudhary, "Digitalizing cultural heritage through metaverse applications: challenges, opportunities, and strategies," vol. 12, no. 1, p. 295, 2024.

[57] S. Münster et al., "Artificial intelligence for digital heritage innovation: Setting up a r&d agenda for europe," vol. 7, no. 2, pp. 794-816, 2024. DOI: https://doi.org/10.3390/heritage7020038

[58] H. T. A. Eyadah and A. A. J. H. Odaibat, "A Forward-Looking Vision to Employ Artificial Intelligence to Preserve Cultural Heritage," vol. 12, no. 5, pp. 109-114, 2024. DOI: https://doi.org/10.11648/j.hss.20241205.12

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).